Essays on International Trade, Growth and Finance

by

Marc-Andreas Muendler

GRAD (University of Munich, Germany) 1998
M.A. (Boston University) 1997

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Economics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Maurice Obstfeld, Co-Chair
Professor David H. Romer, Co-Chair
Professor Daniel L. McFadden
Professor Ann E. Harrison

Spring 2002

The dissertation of Marc-Andreas Muendler is approved:

Co-Chair Date

Co-Chair Date

Date

Date

University of California, Berkeley

Spring 2002

Essays on International Trade, Growth and Finance

Abstract

Essays on International Trade, Growth and Finance

by

Marc-Andreas Muendler

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Maurice Obstfeld, Co-Chair

Professor David H. Romer, Co-Chair

Two concerns in international economics motivate the essays.

I. Does foreign trade harm or foster growth? Two essays look at this question from different perspectives. The first essay takes a dynamic general-equilibrium approach. Contrary to earlier partial-equilibrium models, the essay shows that trade can contribute to reducing the productivity gap between less developed and more advanced regions even if the advanced region hosts most of the innovative industries with dynamic externalities. Productivity may diverge in some cases. Even then both regions generally benefit more from trade than they lose.

The second essay investigates microeconomic effects empirically. It analyzes the channels through which trade has induced productivity change in Brazil after the country liberalized its tariff act in 1990. The facilitated access to foreign inputs plays a minor role for productivity change. However, foreign competition pushes firms to raise efficiency markedly. Counterfactual simulations indicate that this competitive push is a salient source of immediate productivity change. In addition, the shutdown probability of inefficient firms rises with competition from abroad and exerts a positive

impact on aggregate productivity over time.

II. What role does information play in financial markets? Evidence from financial crises suggests that investors possess information about troubled assets early on but do not act upon the information until a crisis looms. This behavior has consequences for the timing and prevention of crises. The two essays in this part introduce an integrated model of information acquisition and portfolio choice. The essays provide new tools for the analysis of information in financial markets, resolve a long-known no-equilibrium paradox, and clear the way for subsequent applied research into international financial crises.

Employing different conjugate prior distributions, the essays demonstrate when investors value information and act on it. More information allows investors to select less risky portfolios. When the asset price is fully revealing, markets do provide information but less than socially desirable. However, more information has a negative effect when becoming commonly known. Commonly shared information moves the asset price closer to the individually expected return, thus reducing the value of the risky asset.

Professor Maurice Obstfeld, Co-Chair

Professor David H. Romer, Co-Chair

To Beatriz

Contents

List of Figures

List of Tables

List of Abbreviations

2SLS	Two-stage least squares
BIS	Bank of International Settlements
BRL	Brazilian Real (Reais)
CARA	Constant absolute risk aversion
EU	European Union
FDI	Foreign direct investment
FE	Fixed effect
HOV	Heckscher-Ohlin-Vanek
IBGE	*Instituto Brasileiro de Geografia e Estatística*
IMF	International Monetary Fund
IV	Instrumental variable
LDC	Less developed country
MNL	Multinomial logit
NLLS	Non-linear least squares
OLS	Ordinary least squares
OP	Olley-Pakes
PIA	*Pesquisa Industrial Anual*
RICE	Rational information choice equilibrium
REE	Rational expectations equilibrium
SECEX	*Secretaria de Comércio Exterior*
SIC	Standard Industrial Classification
TFP	Total factor productivity
US	United States
USD	US Dollar

Acknowledgements

I am indebted and deeply grateful to my advisors Maury Obstfeld, David Romer and Dan McFadden for their guidance and support. My dissertation would have taken a different shape had they not shared their thoughtful comments and encouraged my research throughout my time at Berkeley. While working on the four core chapters of this dissertation, I have benefited from the advice and feedback of many more.

I wish to thank Pranab Bardhan, George Akerlof, Brad DeLong, Chris Shannon, Ethan Ligon, and participants at the Berkeley Development Workshop and LACEA 2000 for many helpful comments and suggestions on chapter 2. Lukman Winoto introduced me to the use of mathematical software.

Ann Harrison, Pranab Bardhan, Andrew Bernard, Aviv Nevo, George Akerlof, Gustavo Gonzaga, Jim Powell, Jim Rauch, Marc Melitz, Mark Machina, Maurício Mesquita, Paul Ruud, and Tom Rothenberg as well as Pete Adams, Pinar Karaca, Tiago Ribeiro and Till von Wachter shared important suggestions on chapter 3, and so did seminar participants at IBGE, PUC-Rio, IPEA Brasília, UC Berkeley, the IDB and the World Bank, UCL, Dartmouth College, U Michigan, JHU, UCSD and LACEA 2001. I thank Gustavo Gonzaga and Humberto Moreira for their help and warm hospitality during my visit to Pontifícia Universidade Católica Rio de Janeiro (PUC-Rio) in 2000 and 2001. I thank Wasmália Bivar, Alexandre Brandão and the team at IBGE's Departamento de Indústria for their patience in providing access to the data on Brazilian manufacturers. Dieter Simons of Siemens S.A. in São Paulo introduced me to the practice of accounting in Brazil. The deflator and capital stock series owe their existence to Adriana Schor and her knowledge of Brazilian data sources and accounting law. Financial support from the Social Science Research Council and the American Council of Learned Societies with an International Predissertation Fellowship (funds from the Ford Foundation), financial support from the World Bank, and mini-grants from IBER and CIDER at UC Berkeley are gratefully acknowledged.

Andy Rose, Barry Eichengreen, Bob Anderson, Chris Shannon and Tom Rothenberg helped me with many insightful remarks on chapters 4 and 5. John Eatwell's views and remarks on international finance during an SSRC Summer Workshop in Applied Economics 1999 inspired me to think about the issue, though the conclusions may differ. Achim Wambach, Sven Rady, and participants at the Munich Research Seminar in July 2000 made important suggestions on an early version of chapter 4. Seminar participants at IBMEC Rio de Janeiro and UC Berkeley shared helpful thoughts on a previous version of chapter 5.

Needless to say, any remaining errors are mine.

Chapter 1

Seizing the Chances of Globalization and Averting its Risks

How can economies benefit from globalization? How can less developed countries engage global markets on their own terms? Recent policy reforms in many countries and the accelerated pace of globalization have brought two concerns to public attention. *Does foreign trade harm or foster growth?* and *What role does information play in financial crises?*.

Does foreign trade harm or foster growth? is the question behind part I of this dissertation. The two chapters of this part investigate the topic under two distinct perspectives—one being purely theoretical and the other mostly empirical in nature. *What role does information play in financial markets?* is the concern of part II. To address the topic, a new theoretical model for information acquisition in financial market is introduced. The two chapters of part II explore the framework under alternative assumptions and lay the ground for subsequent applied research into the question.

Together, the essays aim to provide evidence of how developing countries may seize the chances of globalization and how they can avert some of its risks.

1.1 International trade and growth

The first concern—how foreign trade affects growth—relates to the real side of globalization and the consequences for productivity change in an unequal world. Economically, it touches the realms of growth theory and industrial organization beyond international trade. Economists broadly agree on the welfare benefits of trade in a static world. Both consumers and firms benefit from specialization and from the access to a broader variety of goods on the world market. However, a theoretical debate has ensued and questions the advantage of trade for less developed countries in a dynamic world. Less developed countries may suffer slower growth, it is argued in partial-equilibrium models, if they have to specialize in industries where the learning potential is largely exhausted. Chapter 2, the first chapter in part I, revisits the partial-equilibrium conjectures in a general-equilibrium model. Contrary to those models, the chapter shows theoretically that trade can close the productivity gap between less developed and more advanced regions even if the advanced region hosts most of the innovative industries that generate dynamic externalities. The reason is that monopolistic competition in the innovative industries distorts the wage rate and works to attract modern firms back to the South where labor costs are lower. Yet, productivity between South and North may diverge in some cases. Chapter 2 sets the repeated static gains against dynamic losses for that case and states that even then both regions can benefit more from trade than they lose.

The precise channels of trade effects on growth remain little understood. Chapter 3, the second chapter in part I, sets out to investigate the channels of trade effects empirically. The Brazilian trade liberalization in the early 1990s is used to trace effects of international trade on productivity change. Brazil drastically reduced tariffs and tore down non-tariff barriers for imports between 1990 and 1993. This empirical approach aims at establishing causal relationships between trade and growth. The approach limits the investigation to the short- to medium-term effects of trade reform and holds a magnifying glass to a large sample of manufacturing firms.[1] Three channels through which trade reform affects productivity can be distinguished in the data: (1) Easier access to foreign equipment and materials may allow for a *Foreign Input Push* at the firm level. (2) In the product market, foreign imports and the threat of imports constitute a *Competitive Push* on individual firms. Theory predicts that managers choose to innovate processes and remove slack under fiercer competition. (3) Competition in the product market may also induce more exits and cause a *Competitive Elimination* of inefficient firms.

A newly constructed panel dataset of Brazilian manufacturers contains observations of foreign inputs at the firm level and provides a rare opportunity to separate the first channel. Surprisingly, the suspected *Foreign Input Push* is found to be relatively unimportant. An efficiency difference between foreign and domestic inputs does exist in several sectors but has only a minor impact on productivity. Trade liberalization induces high competitive pressure. The *Competitive Push* on firms to raise their efficiency proves to be a dominant source of productivity change among Brazilian manufacturers in the 1990s. Small changes in the tariff act induce impressive efficiency improvements among surviving firms. When trade barriers fall, the *Competitive Elimination* of the least efficient firms strikes more fiercely. Survival probabilities drop markedly and low-efficiency firms go out of business more frequently. However, simulations indicate that *Competitive Elimination* only exerts an effect on aggregate productivity over time, whereas the *Competitive Push* shows an immediate and strong impact.

[1] A description of the data is relegated to appendix A. For the first time, the study in chapter 3 employs an unbalanced panel of 9,500 medium-sized to large Brazilian manufacturers between 1986 and 1998. Special variables to trace industry turnover and observations of foreign inputs permit refinements in the estimation technique that were not feasible with previous data.

1.2 Information in financial markets

Recent evidence motivates the second concern of this dissertation—what role information may play in financial crises. However, a careful assessment proves to require new theoretical ground.

Experience from the East Asian financial crisis suggests that investors may possess information about troubled securities early on but do not act upon their information until a crisis looms. Why did the first attacks on East Asian currencies only occur in mid 1997 and not before? As early as 1995 the Bank of International Settlements had warned in reports that excessive domestic credit expansion in East Asian economies exposed investors to high risks. Even the International Monetary Fund included subdued warnings for Korea's financial sector in its annual report in 1995 and for Indonesia in 1996. So, the information about financial weaknesses was publicly available early on. Most likely, large investors had similar information even earlier.

Investors with good information on a country's securities may have incentives to conceal private information and to hold their positions until shortly before an expected collapse. If present, this behavior has consequences for the timing and for the prevention of crises. However, theoretical obstacles impede an assessment, especially in the benchmark case of fully revealing asset prices. Most prominently, Grossman and Stiglitz (1980) state in a famous paradox that no rational expectations equilibrium can exist in the benchmark case .

The two chapters in part II employ the theory of conjugate prior distributions to introduce an integrated model of information acquisition and portfolio choice. The model relates to finance theory beyond international finance. Previous models mostly sort investors into informed or uninformed ones. In the present model, investors have more than a binary choice. They can acquire a number of signals and have clear individual incentives for information acquisition. The approach turns information into a public good and makes the analysis amenable to standard economic tools. Even under fully revealing prices, a rational expectations equilibrium does exist both in the securities market and in the information market. The reason is that more information reduces the *ex ante* variance of the asset return from the point of view of the individual investor. Information allows investors to make a more educated portfolio choice and reduces the expected variance of future consumption. Consequently, more information increases the *ex ante* utility of risk averse investors. This gives every investor an individual incentive to acquire information or to remain uninformed if information is too costly. A comprehensive equilibrium results. It clears both the financial market and the market for information.

Irrespective of the varying assumptions in chapters 4 and 5, a risk averse investor likes information in principle. Information sharpens her knowledge, and this allows her to make a more adequate portfolio choice. However, information also has a bad side in financial markets. If information is widely disseminated, the expected excess return of an asset over its opportunity cost falls. In general, prices play a double role: They reflect the opportunity cost of an asset, and they aggregate and disseminate information to everybody. It is this double role of common information that can harm investors in financial markets. If more information gets to the market, this information is at least partly transmitted through price. But then, when rational investors update their information, their expectation of the dividend gets closer to market expectations. In other words, the excess return of an asset over its opportunity cost falls with more information. So, information works to reduce the value of the risky asset when becoming common through price. This is a novel and key insight of the theoretical framework in part II. Investors may not want to act on information: Information reduces the value of a risky security when becoming commonly known through investors' market actions.

Applying the general framework, chapter 4 considers a gamma distributed asset return, Poisson distributed signals (information) and fully revealing prices. It shows that investors buy a positive amount of information whenever markets are large enough, when investors are sufficiently risk averse, or when the variance of the risky asset is relatively high compared to its payoff. Under converse conditions, investors may not want to acquire any information on their own. Chapter 5 explores the general framework in several further directions. Information is not a good in its own right. It is only valuable inasmuch investors anticipate to act upon it. Therefore, risk neutral investors never want to buy information irrespective of the distribution of asset returns and signals (information). With a normally distributed asset return and under fully revealing prices, the negative effect of commonly known information becomes so strong that no investor ever wants to buy any information. However, choosing the normal distribution for the asset return can be an attractive assumption when modeling the more complicated case of a partly informative price, in which some external noise remains. Then, investors do want to acquire information.

When asset prices are noisy, they can only partly reveal other investors' information. As a result, the negative effect of commonly disseminated information is mitigated. Some investors start acquiring information as long as markets are small so that asset price reveals little information to others. When informed investors (news watchers) acquire information, they inflict a negative externality on less informed investors (price watchers) who do not purchase their own information but merely observe the price. The reason is that price watchers rationally anticipate the arrival of information in the market, simultaneously update their beliefs in the same direction and thus make asset price move closer to their own (and average) beliefs. Therefore, the beneficial effect of more precise information never outweighs the loss from a reduced expected return for price watchers.

It remains a project for future research to apply the baseline framework of part II to the specific setting of a financial crisis. However, the models support the hypothesis that investors who possess information may not want to act upon their information immediately. By acting, they would move market price closer to market beliefs and their own beliefs and reduce the asset value. On the other hand, when delaying the adjustment in their portfolio, they also risk a lower return. This tradeoff may help explain the puzzling timing of crises.

Part I

International Trade and Growth

Chapter 2

Trade and growth revisited: Managing to converge, agreeing to diverge

If a pure trade theorist were to advise a less developed country about whether and to what extent it should open up to free trade, she would have to reconcile a large and partly contradictory array of results. Ricardian or Heckscher-Ohlin-Vanek (HOV) models mandate trade liberalization unconditionally. Open up to free trade, the trade theorist would conclude, no matter what your production technologies or factor endowments look like, world markets will start to work so that your comparatively more efficient or endowment-intensive sectors will become export industries and your economy will be better off in the aggregate. An advisor who got to admire new trade theories would be inclined to argue: Irrespective of what the rest of your economy does, if consumers or firms benefit from added varieties of goods, open up to trade and your economy will be better off because consumers and firms benefit from the choice. Here, things already become difficult because the location of industries can be indeterminate, but may matter. Finally, an advisor who adheres to new growth theory will warn: Be careful. If your industries are likely to specialize in low-growth sectors, you may be worse off after liberalization. If you cannot rapidly implement knowledge that is created by other means than learning by doing, or elsewhere, you may become locked into low-tech production and that forever. After all, the advisor won't know.

This chapter sets out to present a simple but comprehensive theoretical framework. The model allows for the four sources of specialization that a trade theorist such as the one above has in mind: international productivity gaps, differences in factor endowments, benefits from variety, and dynamic externalities from knowledge creation. The location of firms is determined endogenously. By construction, it is a worst-case model for a less developed country (LDC). Above all, learning-by-doing externalities will be the only source of productivity growth so that a less advanced region can suffer dynamic losses from trade as argued so often in the past decade. The model is kept simple by assuming explicit functional forms that will give rise to close-form solutions. It may not seem insightful at first to model so much. As wisdom has it, our understanding is sharpened when we isolate effects instead of mixing them. However, once we want to understand the strength of some causes as compared to others, a more comprehensive approach is key.

So far, researchers mostly argued that diverging growth rates would result when dynamic externalities are present in factor accumulation or productivity change. This need not be the case. Eicher (1999) shows in a setting of human capital accumulation that convergence in growth rates (β-convergence) may in fact come about. Similarly, Goh and Olivier (2002) show that convergence may occur when capital goods are traded. Using simple closed-form solutions, I argue in this chapter that convergence can arise in many models of trade and endogenous growth under imperfect competition. The reason is that monopolistic competition distorts the old-style specialization forces. This effect is overlooked in partial equilibrium approaches. The model shares several features with Matsuyama's (1992) and Peletier's (1998) two-sector economies but addresses different questions. Beyond the analysis of an open economy's growth path, the present chapter focuses on the evolution of the international productivity gap and on trade forces that affect it.

In the present general-equilibrium framework, an explicit welfare analysis can be added to growth theory. While it is convenient and mostly correct for closed economies to assert that higher growth means faster welfare increases, the relationship is different for open economies and worth keeping in mind. Open economies benefit from an improvement in their terms of trade when growing more slowly than their trading partners. In addition, repeated static gains from free trade can sum up to vast dynamic gains and outweigh dynamic losses from slower growth for wide ranges of parameters.

Recently, the impact of trade on growth has been reassessed under the auspices of endogenous trade theory and regional economics. Endogenous growth theory seems to make globalization little desirable for LDCs. Young (1991), Stokey (1991), and Peletier (1998) show that trade liberalization may inhibit learning by doing and knowledge creation in LDCs. The reason is that liberalization could induce LDCs to specialize in product lines where the learning potential has been largely exhausted. Xie (1999) shows for a Leontief production technology with intermediate inputs that there can be several, partly offsetting effects of trade on growth. Depending on the relative strength of the forward and backward linkages, trade may harm or spur growth.

A line of argument in regional economics stresses that innovative industries with economies of scale tend to cluster in few locations in order to exploit the increasing returns. Krugman and Venables (1995) argue that, when transportation costs and tariffs fall, manufacturers relocate to a core region where initial demand happens to be high. A periphery will evolve and suffer income losses. This effect can be aggravated when innovation is endogenous (Martin and Ottaviano 1999), but can be partly offset by immobile labor because wages will differ across regions (Puga 1999). Similarly, Matsuyama (1996) shows how a world divided in rich and poor evolves when there are agglomeration effects and countries trade.

The chapter is divided in five sections. Section 2.1 spells out the model, and section 2.2 derives the unique autarky and the unique free-trade equilibrium. Section 2.3 analyzes the dynamics of the 'global economy' and the technology gap between rich and poor regions. Section 2.4 investigates under what conditions free trade can be desirable for a less developed country that has to specialize in low-growth sectors. Section 2.5 concludes.

2.1 The Model

There are two regions called 'North' and 'South' for simplicity. Both regions employ two homogeneous factors of production, capital and skilled labor. Labor is assumed to be perfectly mobile within one region but immobile across borders.Capital is taken to be internationally immobile, too, in order to focus on pure effects of commodity trade. There are two sectors in each region, one 'traditional' and one 'modern' sector. For convenience, call the traditional sector agriculture.This sector makes relatively intensive use of capital (or land). The second sector is manufacturing. Manufacturers employ skilled labor more intensively. They heavily rely on engineering services and software creation, say. All productivity growth stems from the latter sector. The idea is that workers in manufacturing are learning by doing. Their knowledge then benefits the entire economy, as workers can freely change their employment within a region. In agriculture, however, these learning-by-doing effects are largely exhausted.

The economies of each region are endowed with fixed amounts of labor L^i and capital K^i.[1] Consumers are the same everywhere. Their preferences are homothetic. Demand for the agricultural good is standard, but consumers care about varieties in their demand for the 'modern' goods. At every income level, they prefer adding another variety to consuming more of the same varieties.

2.1.1 Production

Let North and South be denoted by $i = N, S$. Then the agricultural sector in region i produces X^i with a Cobb-Douglas technology at time t:

$$X^i(t) = \left[A^i(t)L^i_X(t)\right]^\gamma \left[K^i(t)\right]^{1-\gamma}, \qquad \gamma \in (0,1). \tag{2.1}$$

[1] Allowing for capital accumulation does not change in the main results.

The variable A^i denotes the economy-wide labor productivity. L^i_X is the number of region i's workers employed in sector X^i. The product of labor with its efficiency $A^i L^i$ can also be thought of as a stock of skills or human capital, justifying the assumption that labor here means skilled labor. K^i denotes capital employed in the agricultural sector. It does not carry a subscript because the modern sector will not employ capital.

The modern sector, on the other hand, looks like Krugman's (1980) one-sector economy. It consists of a measure of N^i firms. (N^i will be determined endogenously in equilibrium.) Each single firm n manufactures a quantity z^i_n of goods under an identical increasing-returns-to-scale production technology that uses skilled labor as its only input:

$$z^i_n(t) = A^i(t) \left[L^i_n(t) - L_0 \right]. \tag{2.2}$$

L^i_n denotes the number of workers employed in region i's firm n, and L_0 is a fixed amount of labor that has to be employed each period to keep the firm running. For simplicity, L_0 is the same in both regions and it is not sunk. So, the increasing returns to scale are never exhausted in the modern sector. While natural monopolies can lose their economies of scale over time, there will always be new entrants and innovators that again exhibit scale economies for some period of time.

The above production technologies embody two classical and one modern source of trade specialization. First, since labor productivity A^i may differ between North and South and $\gamma < 1$, Ricardian trade theory predicts that the region with the higher labor productivity A^i specializes in modern goods production, all else equal. Second, HOV theory predicts that the region with the higher capital-labor ratio will *ceteris paribus* specialize in agriculture. HOV theory also predicts that the specialization after trade will be incomplete when the two regions are sufficiently similar. Third, new trade theory predicts that both regions will engage in intraindustry trade of manufactured goods. That is, both regions will produce varieties of the modern good and consume all foreign varieties along with the domestic varieties. Due to increasing returns to scale, monopolistic competition will arise in the modern sector and prices will remain above marginal cost. However, freely entering firms will compete away all rents. The only benefit from hosting the modern sector within the own borders stems from a dynamic externality in technological change. This component gives rise to the primary concern here: Does trade hurt the South?

2.1.2 Technological change

Workers employed in manufacturing learn from every unit they manufacture. However, modern firms do not internalize this knowledge creation. It is a byproduct of their manufacturing activity and, as such, a 'dynamic externality'. Similar externalities were elaborated in Arrow (1962) and, more recently, P. Romer (1986). Many forms of endogenous growth stem from sources that cannot be internalized completely by markets because knowledge is a public good so that its creation is generally underpriced. Under the assumption that there will be a continuum of modern firms, each producing exactly one variety, knowledge creation can be given the following form of a pure externality:

$$\dot{A}^i(t) \equiv B \int_0^{N^i} z_n^i(t)\,\mathrm{d}n, \tag{2.3}$$

where B is some positive constant and identical in both regions.

In a more realistic model, learning by doing in agriculture would also contribute to this knowledge creation. However, employees in the modern sector accumulate skills more rapidly, whereas the learning by doing potential is largely exhausted in agriculture. Relaxing the assumption and explicitly including knowledge creation in agriculture would not change the main results of the model as long as learning by doing is faster in industry.

2.1.3 Demand

Consumers are identical in both regions. Their preferences take the form that Dixit and Norman (1980) introduced for simultaneous interindustry and intraindustry trade. Consider the following consumption index of modern goods and the related price index

$$D^i \equiv \left(\int_0^N (d_n^i)^\alpha \, \mathrm{d}n\right)^{\frac{1}{\alpha}} \quad \text{and} \tag{2.4}$$

$$P \equiv \left(\int_0^N (p_n)^{-\frac{\alpha}{1-\alpha}} \, \mathrm{d}n\right)^{-\frac{1-\alpha}{\alpha}}, \tag{2.5}$$

which are harmonic means of the N consumed varieties and their prices. A representative consumer in region i has an instantaneous utility u

$$u(C^i, D^i) = (C^i)^{1-\theta}(D^i)^\theta = (C^i)^{1-\theta}\left(\int_0^N (d_n^i)^\alpha \, \mathrm{d}n\right)^{\frac{\theta}{\alpha}} \tag{2.6}$$

of consuming a quantity C^i of the agricultural good, and quantities d_n^i of each variety n of the modern good. The representative consumer purchases a measure N of these modern goods. The parameters α and θ are both restricted to values between zero and one: $\alpha, \theta \in (0, 1)$.[2] In every period, each household maximizes (2.6) with respect to C^i, d_n^i, and N, such that the budget constraint $C^i + \int_0^N p_n d_n^i \, \mathrm{d}n \leq Y^i$ is satisfied. Here and from now on the agricultural good is the *numéraire* with $P_X \equiv 1$, while p_n is the unit price of variety n of the modern good. For the consumer's problem to be well-defined, the constraint $N \leq \bar{N}$ must be satisfied in addition to the budget constraint. $\bar{N}$ is the total number of varieties available to the consumer.

Then, the resulting demand for the agricultural good and each variety of the modern good become

$$C^i = (1-\theta)\, Y^i \tag{2.7}$$

and

$$d_n^i = \theta Y^i \left(\frac{P^\alpha}{p_n}\right)^{\frac{1}{1-\alpha}}, \tag{2.8}$$

[2] θ has to lie between zero and one for $u(\cdot)$ to be a well-defined utility function. The requirement that α not exceed unity can be justified from the implied elasticities of substitution. In a Cobb-Douglas utility function, the elasticity of substitution between C and a modern good d_j is one. However, the elasticity of substitution between one modern good and another modern good is $1/(1-\alpha)$. In order to obtain stronger substitutability among modern goods than between them and the agricultural product, $1/(1-\alpha) \in (1, \infty)$ is needed, i.e. $\alpha \in (0, 1)$.

respectively. The price elasticity of demand for a modern good n is

$$\varepsilon_{d_n,p_n} = -\frac{1}{1-\alpha}\left[1-\alpha\left(\frac{P}{p_n}\right)^{\frac{\alpha}{1-\alpha}}\right]. \tag{2.9}$$

For the above utility function, consumers prefer adding another variety to consuming more quantity of each variety. That is, when modern goods sell at sufficiently close prices, the optimal N equals $\bar{N}$ (or is zero). As long as $N < \bar{N}$, consumers lower the quantity d_n^i (for all the $n \in [0, N]$ they are consuming) and add another variety (to increase N), while they still satisfy the budget constraint.

Given these demand functions, indirect utility at each instant τ becomes

$$u(\tau) = T\, Y^i(\tau)\, P(\tau)^{-\theta}, \tag{2.10}$$

where $T \equiv (1-\theta)^{1-\theta}\,\theta^{\theta}$.

2.2 Autarky and Free Trade Equilibrium

Since only labor is employed in both sectors of industry, the entire per period equilibrium allocation and all prices can be expressed in terms of the *share of the labor force employed in the modern sector*. Call this share λ^i:

$$\lambda^i(t) \equiv \frac{\int_0^{N^i} L_n^i(t)\,\mathrm{d}n}{L^i} = \frac{L_N^i(t)}{L^i} \in [0,1], \tag{2.11}$$

where $L_N^i(t) \equiv \int_0^{N^i} L_n^i(t)\,\mathrm{d}n$. Ultimately, the equilibrium growth rate will also be determined completely by this labor share.

In this section, two convenient equilibrium relationships are derived first: the equilibrium scale of production of modern firms and the equilibrium number of modern firms. They take the same functional form under autarky and free trade. Then, the autarky equilibrium and finally the world trade equilibrium will be derived. For ease of notation, the time variable is dropped for parts of the exposition with the understanding that all endogenous variables remain time dependent and that the per period equilibrium values of the variables may change over time.

2.2.1 Monopolistic competition in the modern sector

In order to enter the market for a new variety, a modern firm must incur the fixed cost $w^i L_0$, where w^i is the wage rate in region i. Since there are no economies of scope or sunk cost, incumbent firms have no advantage over entrants. Hence, one can assume without loss of generality that each firm in the modern sector can manufacture only one variety. Under increasing returns to scale, no second firm can successfully compete in the market for any single variety. Each variety is thus manufactured by one and only one firm. However, free entry into neighboring markets for modern goods will drive profits down to zero. Given production technology (2.2), each firm's cost function is $C(w^i, z_n^i) = w^i z_n^i / A^i + w^i L_0$. A firm n finds it optimal to employ $L_n^i = z_n^i / A^i + L_0$ workers for the production of a positive quantity z_n^i of variety n (for $z_n^i = 0$, optimal labor demand is $L_n^i = 0$ because L_0 is not sunk). Note that every firm needs the fixed amount of labor for its operation in each period. The fixed factor is employed again and again, as long as the firm remains in business. Since every firm is a monopolist in the market for its own variety, it will take into account how demand responds to its supply decision. So, the optimal quantity z_n^i is determined by the profit maximizing condition that marginal revenue equal marginal cost $p_n^i \, (1 - \epsilon_{p_n, d_n}) = w^i / A^i$.

Neglecting equilibria in which varieties are sold at different prices, one can follow Dixit and Norman (1980) and Krugman (1980) and assume that the equilibrium is symmetric. This simplifies the analysis considerably. Let prices for modern goods p_n^i satisfy $p_n^i = p^i \, \forall n$. Then, for a sufficiently high number of firms in the modern sector, each firm will set its quantity so that consumers have to pay the price

$$p^i \simeq \frac{1}{\alpha} \frac{w^i}{A^i}, \tag{2.12}$$

where the mark-up $1/\alpha$ approximately derives from demand elasticity (2.9) for a large measure of firms. The approximation $1 - \epsilon_{p_n, d_n} \approx 1 - 1/\epsilon_{d_n, p_n} \approx \alpha$ only implies that firms cannot squeeze out the entire consumer rent they would optimally choose to extract. Thus, the resulting number of entrants will be lower than it could be if firms were allowed to take the term into account. However, the allocation of labor to the modern sector, λ^i, will not depend on this simplification.

In an unregulated market, entry will occur until profits are driven down to zero: $\pi_n^i = \left(p_n^i - w^i/A^i\right) z_n^i - w^i L_0 = 0$. Using the quantity decision implied by (2.12), each firm will produce at the break-even scale

$$z^i = \alpha \cdot A^i \frac{L_0}{1-\alpha} \tag{2.13}$$

and employ $L^i = L_0/(1-\alpha)$ workers in a symmetric equilibrium. Were firms not able to charge a premium over marginal cost, they could not sustain production because their fixed cost would not be covered. The quantity choice that results from monopolistic competition is, as (2.13) shows, lower by a factor of α than it would be in a social optimum (where a social planner would need to compensate firms for their fixed cost through a lump sum transfer).

2.2.2 Equilibrium number of varieties

In general equilibrium, the number of modern firms will be determined by the relative size of the manufacturing sector. Solving for the equilibrium levels of variables turns out to be much easier when looking at the economy from the income side. The modern sector exclusively employs labor. It follows immediately from (2.13) that each manufacturer generates revenues of $p^i z^i = p^i \cdot [\alpha/(1-\alpha)]A^i L_0$. Since monopolistic competition drives profits down to zero, all these revenues must go to workers. Thus,

$$w^i \cdot \lambda^i L^i = N^i \cdot p^i \frac{\alpha}{1-\alpha} A^i L_0 \tag{2.14}$$

in the modern sector.

Together with the monopolistic pricing rule (2.12), this income identity yields a simple relationship between the number of firms N^i and the equilibrium labor share λ^i:

$$N^i = (1-\alpha)\frac{L^i}{L_0} \cdot \lambda^i. \tag{2.15}$$

The smaller the fixed amount of labor L_0 or the higher the monopoly power of firms (the lower α), the more firms enter. Independent of the concrete parameter values, entering firms will compete all rents away.

2.2.3 Autarky equilibrium

Four markets have to clear in region i in autarky. The labor market, the capital market, and the two commodity markets. Take the two commodity markets first. Since prices for all varieties are equal in symmetric equilibrium ($p_n = p$), demand for each variety (2.8) simplifies to $d_n^i = d^i = \theta Y^i / N^i p^i$. So, the market clearing condition for each variety becomes:

$$d^i = \frac{\theta Y^i}{N^i p^i} \; = \alpha A^i \frac{L_0}{1-\alpha} = z^i. \tag{2.16}$$

Similarly, the agricultural goods market clears if

$$C^i = (1-\theta) Y^i \; = \; X^i. \tag{2.17}$$

By expressing (2.16) and (2.17) in terms of λ^i, labor market clearing was implicitly imposed: $L_N^i + L_X^i = L^i$. The last among the four markets—the capital market—must clear by Walras' Law.

The interest rate r^i is such that the agricultural sector finds it optimal to employ all supplied capital. The labor market clears at a wage rate w^i equal to the marginal product in both sectors, the market for modern goods clears at a price p^i given by (2.12), and the agricultural good sells at a price of $P_X = 1$. So,

$$w^i = \frac{\gamma X^i}{(1-\lambda^i) L^i}, \tag{2.18}$$

$$r^i = \frac{(1-\gamma) X^i}{K^i}, \tag{2.19}$$

$$p^i = \frac{1}{\alpha} \frac{w^i}{A^i}. \tag{2.20}$$

Agricultural production X^i is a function of the labor share λ^i, the capital stock and parameters, $X_i = [A^i (1-\lambda^i) L^i]^\gamma [K^i]^{1-\gamma}$.

Hence, all equilibrium prices and quantities in a period can be expressed as functions of the labor share λ^i. What is the equilibrium labor share λ^i? Total income must equal total consumption expenditure in equilibrium

$$w^i L^i + r^i K^i = (1-\theta) Y^i + \theta Y^i = Y^i. \tag{2.21}$$

Using this income and expenditure relationship (2.21) along with the market clearing and price equations (2.16) through (2.20) yields the equilibrium. There are six equations in six unknowns λ^i, N^i, p^i, w^i, r^i, Y^i. The following statement summarizes what the equilibrium looks like.

Proposition 2.1 *In autarky, the equilibrium share of workers employed in agriculture is*

$$1 - \lambda^i_{aut} = \frac{\gamma(1-\theta)}{\theta + \gamma(1-\theta)}. \tag{2.22}$$

The size of the modern sector can be expressed with the equilibrium number of modern firms

$$N^i_{aut} = \frac{(1-\alpha)\,\theta}{\theta + \gamma(1-\theta)} \frac{L^i}{L_0}, \tag{2.23}$$

so that productivity grows at a rate

$$g^i_{A,aut} \equiv \alpha B L^i \, \frac{\theta}{\theta + \gamma(1-\theta)} \tag{2.24}$$

in equilibrium.

Proof. The two market clearing conditions, (2.16) and (2.17) in the text, the three price equations (2.18), (2.19), and (2.20), along with the income-expenditure relationship (2.21), constitute a system of six equations in six unknowns λ^i, N^i, p^i, w^i, r^i, Y^i.

To derive the equilibrium, start with market clearing in the modern sector: Using the three price relationships—(2.18), (2.19), (2.20)—in (2.21), income (2.16) can be rewritten as

$$Y^i = N^i p^i z^i = \frac{\gamma X^i}{1-\lambda^i} + (1-\gamma) X^i. \tag{2.25}$$

By (2.15) in the text, the equilibrium number of firms $N^i = (1-\alpha) L^i \lambda^i / L_0$ can be immediately derived from (2.20) and (2.14). Using this, again along with the price for modern goods (2.20),

$N^i p^i z^i$ becomes $N^i p^i z^i = \gamma X^i \lambda^i/(1-\lambda^i)$. Substituting for $N^i p^i z^i$ in (2.25) and solving out for λ^i yields (2.22) in the text. The equilibrium number of firms (2.23) and productivity growth (2.24) follow readily. ■

None of these equilibrium variables changes over time. Thus, the autarky equilibrium is also a steady-state. The allocation of labor to the modern sector increases whenever modern goods are in high demand (large θ), and falls when labor is intensively used in agriculture (large γ). The equilibrium labor allocation is independent of the level of labor skills, A^i, since these skills are equally applicable in both sectors. It is also independent of the elasticity of substitution between modern goods ($1/(1-\alpha)$) since it only matters for the number of modern firms, not for the size of the sector as a whole. In general equilibrium, the number of modern firms is directly proportional to the labor share in the modern sector by (2.15). The total of modern goods is $N^i z^i = \alpha A^i \lambda^i L^i$ by (2.13) and (2.15). Productivity growth stems exclusively from learning by doing in the modern sector. By (2.3), it equals $BN^i z^i$. The learning function thus takes the value $\dot{A}^i = A^i\, \alpha\theta B L^i/(\theta + (1-\theta)\gamma)$ in an autarky equilibrium.

2.2.4 Equilibrium under free trade

Let both regions open up completely to free trade. Call the home region i and the foreign region $-i$. Assume that there are no transport costs or tariffs after trade liberalization. In order to keep results simple, restrict attention to an equilibrium in which all varieties from one region sell at the same price world wide. All Southern goods sell at price p^S and all Northern goods at p^N.

The price relationships and market clearing conditions that applied to autarky continue to hold in a world trade equilibrium—with three exceptions: the market clearing condition for the agricultural good and the market clearing conditions for the Southern and Northern modern goods. Market clearing of the agricultural good (2.17) generalizes to

$$C^i + C^{-i} = (1-\theta)\left(Y^i + Y^{-i}\right) \;=\; X^i + X^{-i}. \tag{2.26}$$

If specialization after trade liberalization is not complete, N^S modern firms will locate in the South and N^N firms will manufacture in the North. Denote Southern consumers' demand for modern goods from region j by $d^{j,S}$. More generally, $d^{j,i}$ modern goods are delivered from region j to consumers in region i. Then, market clearing for modern commodities manufactured in region j requires that $d^{j,i} + d^{j,-i} = z^j$ for $j = S, N$. Since all Southern goods sell at p^S and all Northern goods at p^N, demand (2.8) for modern goods from region j simplifies to

$$d^{j,i} = \frac{\theta Y^i}{\left[N^S (p^S)^{-\frac{\alpha}{1-\alpha}} + N^N (p^N)^{-\frac{\alpha}{1-\alpha}}\right]} \frac{1}{(p^j)^{\frac{1}{1-\alpha}}} \qquad j = S, N \tag{2.27}$$

in region i. Thus, market clearing for goods from region j can be written as

$$\begin{aligned} d^{j,S} + d^{j,N} &= \frac{\theta \left(Y^S + Y^N\right)}{\left[N^S (p^S)^{-\frac{\alpha}{1-\alpha}} + N^N (p^N)^{-\frac{\alpha}{1-\alpha}}\right]} \cdot \frac{1}{(p^j)^{\frac{1}{1-\alpha}}} \\ &= \frac{\alpha}{1-\alpha} L_0 \cdot A^j = z^j \qquad j = S, N. \end{aligned} \tag{2.28}$$

Dividing (2.28) for the North by (2.28) for the South, yields the price ratio in equilibrium

$$\frac{p^S}{p^N} = \left(\frac{A^N}{A^S}\right)^{1-\alpha}. \tag{2.29}$$

In addition to the three market clearing conditions (2.26) and (2.28), labor markets and capital markets must clear in both regions. As in autarky, expressing the equilibrium with labor shares λ^i and λ^{-i} implicitly imposes labor market clearing in both regions. Capital markets must clear in both countries by Walras' Law. Thus, the world trade equilibrium can be described by the price relationships (2.18), (2.19) and (2.20) as in autarky, and the two income-expenditure relationships (2.14) and (2.21), which express income generated in the modern sector and income generated in the entire economy, respectively. Each of these conditions must hold for both regions i and $-i$. Together with market clearing for the agricultural good (2.26) and the modern commodities (2.27) (the latter applied to both region i and $-i$), these relationships constitute a system of thirteen equations in thirteen unknowns λ^i, λ^{-i}, N^i, N^{-i}, p^i, p^{-i}, w^i, w^{-i}, r^i, r^{-i}, Y^i, Y^{-i}, and P_X. The number of equations is odd because there is only one market clearing condition for the agricultural good.

Just as for the derivation of autarky equilibrium, it proves convenient to look at the economy from the income and spending side. World-wide revenues in the modern sector must equal world-wide spending on modern goods,

$$\begin{aligned} p^S N^S z^S + p^N N^N z^N &= \left(p^S N^S A^S + p^N N^N A^N\right) \frac{\alpha}{1-\alpha} L_0 \\ &= \theta \left(Y^S + Y^N\right). \end{aligned} \tag{2.30}$$

For convenience, the two market clearing conditions in (2.28) can be replaced by imposing the implied world price ratio (2.29) and income-expenditure relationship (2.30) instead. The unique world trade equilibrium—in the three equations (2.26), (2.29) and (2.30) along with the ten price and income relationships (2.18), (2.19), (2.20), (2.14), (2.21)—has an intuitive closed form.

To state the solution more succinctly in proposition 2.2, I define two handy variables called *specialization forces*. If country i is relatively abundantly endowed with labor, free trade will *ceteris paribus* cause an expansion in the modern sector. Similarly, if country i is relatively abundantly endowed with capital, its agricultural sector will expand after trade. Let Λ^i denote the specialization force from labor endowments that pushes country i to more agricultural production, and Γ^i denote the specialization force from capital endowments that pushes the same country i to more modern production. These specialization forces from labor endowments and from capital endowments can be defined rigorously as

$$\Lambda^i(t) \equiv 1 + \left(\frac{A^{-i}(t)}{A^i(t)}\right)^{\alpha} \frac{L^{-i}}{L^i} \quad \text{and} \quad \Gamma^i(t) \equiv 1 + \left(\frac{A^{-i}(t)}{A^i(t)}\right)^{\gamma\frac{1-\alpha}{1-\gamma}} \frac{K^{-i}}{K^i}, \tag{2.31}$$

respectively. The term $A^{-i}/A^i(t)$ is the *productivity gap* between the two regions $-i$ and i. The factors $\left(A^{-i}/A^i\right)^{\alpha}$ and $\left(A^{-i}/A^i\right)^{(1-\alpha)\frac{\gamma}{1-\gamma}}$ affect both specialization forces in this particular form due to monopolistic competition in the modern sector. The factors equal the relative factor prices in equilibrium. They are concave or convex functions of the productivity gap A^{-i}/A^i, depending on the relative magnitude of the parameters α and γ. So, α and γ in the powers on A^{-i}/A^i determine the behavior of the specialization forces. Their presence will be the key to growth convergence.

With these definitions, the trade equilibrium can be expressed in the following manner.

Proposition 2.2 *After trade liberalization, the equilibrium share of workers employed in agriculture is*

$$1-\lambda^i_{trade}(t)=\frac{\gamma(1-\theta)}{\theta+\gamma(1-\theta)}\frac{\Lambda^i(t)}{\Gamma^i(t)}, \tag{2.32}$$

The size of the modern sector is given by the equilibrium number of modern firms

$$N^i_{trade}(t)=\frac{1-\alpha}{\theta+\gamma(1-\theta)}\frac{L^i}{L_0}\left(\theta+\gamma(1-\theta)\frac{\Gamma^i(t)-\Lambda^i(t)}{\Gamma^i(t)}\right), \tag{2.33}$$

so that productivity in country i grows at a rate

$$g^i_{A,trade}(t)=\alpha BL^i\left(1-\frac{\gamma(1-\theta)}{\theta+\gamma(1-\theta)}\frac{\Lambda^i(t)}{\Gamma^i(t)}\right). \tag{2.34}$$

The two factor price ratios are

$$\frac{w^{-i}_{trade}}{w^i_{trade}}(t)=\left(\frac{A^{-i}_{trade}(t)}{A^i_{trade}(t)}\right)^{\alpha} \quad \text{and} \quad \frac{r^{-i}_{trade}}{r^i_{trade}}(t)=\left(\frac{A^{-i}_{trade}(t)}{A^i_{trade}(t)}\right)^{\gamma\frac{\gamma}{1-\alpha}}, \tag{2.35}$$

respectively.

Corollary 2.2.1 *If there are no fixed costs in the modern sector ($L_0=0$), or if A^i is treated as total factor productivity in the agricultural sector (or both), then the respective specialization forces and factor price ratios are as in table 2.1.*

Corollary 2.2.2 *For $\frac{\Lambda^i(t)}{\Gamma^i(t)}\geq 1+\frac{\theta}{\gamma(1-\theta)}$, region i completely specializes in agriculture and stops growing.*

Corollary 2.2.3 *Since capital is immobile across regions, each region will host an agricultural sector and cannot specialize completely in the modern sector.*

Proof. The three market clearing conditions, (2.26), (2.29) and (2.30) along with the six $(3\cdot 2)$ price equations (2.18), (2.19), (2.20), and the four $(2\cdot 2)$ income relationships (2.14) and (2.21) constitute an equation system in thirteen equations and thirteen unknowns: λ^i, λ^{-i}, N^i, N^{-i}, p^i, p^{-i}, w^i, w^{-i}, r^i, r^{-i}, Y^i, Y^{-i}, and P_X. One equation is redundant so that the price of the agricultural good can be set to unity.

Using the three price relationships once for market clearing in the modern sector and once for market clearing in agriculture, λ^i can be expressed in terms of agricultural output. That yields

$$1-\lambda^i = \frac{\gamma(1-\theta)}{\theta+\gamma(1-\theta)} \left(1+\left(\frac{A^{-i}}{A^i}\right)^{\delta_L} \frac{L^{-i}}{L^i}\right) \frac{X^i}{X^i+X^{-i}}. \tag{2.36}$$

Relationship (2.36) must also hold for economy $-i$, so that

$$\frac{1-\lambda^i}{1-\lambda^{-i}} = \frac{1+\left(\frac{A^{-i}}{A^i}\right)^{\delta_L} \frac{L^{-i}}{L^i}}{1+\left(\frac{A^i}{A^{-i}}\right)^{\delta_L} \frac{L^i}{L^{-i}}} \frac{X^i}{X^{-i}} = \left(\frac{A^{-i}}{A^i}\right)^{\delta_L} \frac{L^{-i}}{L^i} \frac{X^i}{X^{-i}}. \tag{2.37}$$

By (2.1),

$$\frac{X^i}{X^{-i}} = \left(\frac{A^i}{A^{-i}}\right)^{\delta_A} \left(\frac{1-\lambda^i}{1-\lambda^{-i}}\right)^{\gamma} \left(\frac{L^i}{L^{-i}}\right)^{\gamma} \left(\frac{K^i}{K^{-i}}\right)^{1-\gamma}, \tag{2.38}$$

where $\delta \in \{\gamma, 1\}$. Using (2.38) in (2.37) and solving out for $\frac{1-\lambda^i}{1-\lambda^{-i}}$ yields

$$\frac{1-\lambda^i}{1-\lambda^{-i}} = \left(\frac{A^{-i}}{A^i}\right)^{\frac{\delta_L-\delta_A}{1-\gamma}} \frac{L^{-i}}{L^i} \frac{K^i}{K^{-i}}.$$

Using this in (2.38) again yields

$$\frac{X^{-i}}{X^i} = \left(\frac{A^{-i}}{A^i}\right)^{\frac{\delta_A-\gamma\delta_L}{1-\gamma}} \frac{K^{-i}}{K^i}$$

so that, by (2.36),

$$1-\lambda^i = \frac{1+\left(\frac{A^{-i}}{A^i}\right)^{\delta_L} \frac{L^{-i}}{L^i}}{1+\left(\frac{A^{-i}}{A^i}\right)^{\delta_K} \frac{K^{-i}}{K^i}} \frac{\gamma(1-\theta)}{\theta+\gamma(1-\theta)}, \tag{2.39}$$

where $\delta_K \equiv \frac{\delta_A-\gamma\delta_L}{1-\gamma}$. This establishes proposition 2.2 and corollary 2.2.1 for $\delta_A \in \{\gamma, 1\}$ and $\delta_L \in \{\alpha, 1\}$.

For a proof of corollary 2.2.3, suppose that $\lambda^i = 1$. Then capital in region i is unemployed if it cannot flow to region $-i$, and the marginal product of capital is infinite, as is the interest rate. This cannot be an equilibrium. More formally, $\lambda^i = 1$ implies $\frac{\Lambda^i}{\Gamma^i} \leq 0$ by (2.32), which is impossible. Corollary 2.2.2 immediately follows from (2.32) with $\lambda^i = 0$. ■

After trade liberalization, the equilibrium labor share in agriculture differs from the autarky equilibrium by a factor of Λ^i/Γ^i. The higher Λ^i, that is the higher the labor endowment abroad

Table 2.1: Trade Equilibria for Classical and Modern Economies

	Specialization Force		Factor Price Ratios		Elasticities	
Type of Economy	Λ^i	Γ^i	$\frac{w^{-i}}{w^i}$	$\frac{r^{-i}}{r^i}$	δ_L	δ_K
Classical economy						
A^i: L.-Prod.	$1+\frac{A^{-i}}{A^i}\frac{L^{-i}}{L^i}$	$1+\frac{K^{-i}}{K^i}$	$\frac{A^{-i}}{A^i}$	1	1	0
A^i: TFP	$1+\frac{A^{-i}}{A^i}\frac{L^{-i}}{L^i}$	$1+\frac{A^{-i}}{A^i}\frac{K^{-i}}{K^i}$	$\frac{A^{-i}}{A^i}$	$\frac{A^{-i}}{A^i}$	1	1
$(\alpha=1, L_0=0)$	>1	>1			$=1$	≤ 1
Modern economy						
A^i: L.-Prod.	$1+\left(\frac{A^{-i}}{A^i}\right)^{\alpha}\frac{L^{-i}}{L^i}$	$1+\left(\frac{A^{-i}}{A^i}\right)^{\gamma\frac{1-\alpha}{1-\gamma}}\frac{K^{-i}}{K^i}$	$\left(\frac{A^{-i}}{A^i}\right)^{\alpha}$	$\left(\frac{A^{-i}}{A^i}\right)^{\gamma\frac{1-\alpha}{1-\gamma}}$	α	$\gamma\frac{1-\alpha}{1-\gamma} \lesseqgtr 1$
A^i: TFP	$1+\left(\frac{A^{-i}}{A^i}\right)^{\alpha}\frac{L^{-i}}{L^i}$	$1+\left(\frac{A^{-i}}{A^i}\right)^{\frac{1-\alpha\gamma}{1-\gamma}}\frac{K^{-i}}{K^i}$	$\left(\frac{A^{-i}}{A^i}\right)^{\alpha}$	$\left(\frac{A^{-i}}{A^i}\right)^{\frac{1-\alpha\gamma}{1-\gamma}}$	α	$\frac{1-\alpha\gamma}{1-\gamma} > 1$
$(\alpha \in (0,1), L_0>0)$	>1	>1			<1	$\neq 1$

relative to the labor endowment at home, the more workers at home become employed in agriculture after trade. Similarly, the lower Γ^i, the more workers at home become employed in agriculture. The relative specialization forces change over time so that the two regional economies need no longer find themselves in steady states. To reap the full benefits of trade liberalization, factor markets in both regions must be sufficiently flexible and adjust to ongoing economic changes. Note that the specialization forces for regions i and $-i$ are not the inverses of each other. Rather, $\Lambda^i = 1 + 1/(\Lambda^{-i} - 1)$ and $\Gamma^i = 1 + 1/(\Gamma^{-i} - 1)$.

Table 2.1 contrasts the economy mainly under consideration here with related economies. The 'classical economies' have a manufacturing sector with constant returns to scale so that modern output is produced under technology $Z^i = A^i L^i$ (and $L_0 = 0$). The equilibrium number of firms is indeterminate in such an economy, and can be set to $N^i = N^{-i} \equiv 1$ for convenience. The productivity coefficient A^i can be understood as *labor productivity* if agricultural production takes the form $X^i(t) = \left[A^i(t)L^i_X(t)\right]^\gamma \left[K^i(t)\right]^{1-\gamma}$ as in (2.1). It can be interpreted as total factor productivity (TFP) if agricultural production is modified to $X^i(t) = A^i(t) \left[L^i_X(t)\right]^\gamma \left[K^i(t)\right]^{1-\gamma}$.

In the absence of a productivity gap ($A^i = A^{-i}$), factor price equalization obtains as in HOV trade theory. In this sense, the 'classical economy' with A^i being labor productivity seems to be a natural benchmark case. It results in factor price equalization for the interest rate, but not for the wage rate. One could call this 'conditional factor price equalization'—conditional on productivity differences. Empirically, this is a typical pattern. Real interest rates are roughly equal across countries, even between richer and poorer regions, but real wages differ substantially. So, it seems slightly more appropriate in the present context to view A^i to mean labor productivity.

Intraindustry trade ends with simple 'conditional factor price equalization' due to the price distortion from monopolistic competition (which is necessary for modern firms to recover their fixed cost). The international wage differential becomes $\left(A^{-i}/A^i\right)^\alpha < \left(A^{-i}/A^i\right)$. This distortion gives rise to the possibility that convergence across regions can occur even though growth stems from a dynamic externality.

2.3 Managing to Converge: The Technology Gap under Free Trade

How does the specialization force Λ^i/Γ^i and how does the productivity gap between countries A^{-i}/A^i evolve in world trade equilibrium over time? This section will show that the dynamics largely depend on the type of economies that participate in international trade. Regions that strongly engage in intraindustry trade tend to converge, whereas economies that concentrate in classical interindustry trade tend to diverge after trade liberalization. Since productivity growth is proportional to an economy's labor endowment, $g_A^i = \alpha B \lambda^i L^i$, there would be strong autonomous forces for divergence if $L^{-i} \neq L^i$ in this world economy. In order to concentrate on purely endogenous forces of divergence or convergence, set $L^{-i} = L^i = 1$ for the discussion in this section.

Proposition 2.3 *After trade liberalization, the productivity gap A^{-i}/A^i changes at a rate*

$$\begin{aligned} \left(\dot{\overline{\frac{A^{-i}}{A^i}}}\right) \Big/ \left(\frac{A^{-i}}{A^i}\right) &= g_A^{-i} - g_A^i = \frac{\alpha\gamma(1-\theta)B}{\theta+\gamma(1-\theta)}\left(\frac{\Lambda^i}{\Gamma^i} - \frac{\Lambda^{-i}}{\Gamma^{-i}}\right) \\ &= \frac{\alpha\gamma(1-\theta)B}{\theta+\gamma(1-\theta)}\frac{\Lambda^i}{\Gamma^i}\left(1-\left(\frac{A^{-i}}{A^i}\right)^{\delta_K-\delta_L}\frac{K^{-i}}{K^i}\right) \end{aligned} \tag{2.40}$$

for $L^{-i} = L^i$. The coefficients δ_K and δ_L are the elasticities of the factor price ratios with respect to the productivity gap, as given in table 2.1 (p. 25).

Proof. Taking the time-derivative of A^{-i}/A^i and using the equilibrium productivity growth rates (2.34) along with definitions $\Lambda^i \equiv 1 + \left(A^{-i}/A^i\right)^{\delta_L}$ and $\Gamma^i \equiv 1 + \left(A^{-i}/A^i\right)^{\delta_K}\left(K^{-i}/K^i\right)$, yields (2.40). ■

2.3.1 Convergence under free trade

The evolution of the international productivity gap as described in (2.40) allows for rich patterns of divergence or convergence. Divergence in productivity levels will occur if function (2.40) is increasing. Convergence can occur, on the other hand, if (2.40) is decreasing in a neighborhood of

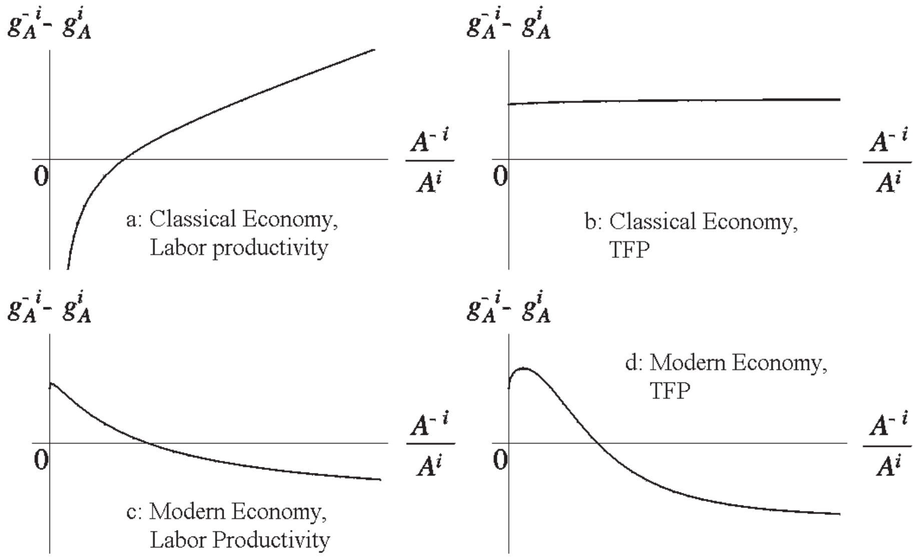


Figure 2.1: Divergence and convergence patterns

some steady-state technology gap A_0^{-i}/A_0^i. Figure 2.1 depicts some examples for the four types of economies in table 2.1 (p. 25).[3] In figure 2.1, the two 'classical economies' are depicted in the upper row. They diverge after trade liberalization, whereas 'modern economies' converge in productivity levels as depicted in the lower row. The examples in figure 2.1 are representative of more general cases to be derived below.

For convergence to occur in a neighborhood of some A_0^{-i}/A_0^i, the right hand side of (2.40) must be decreasing. Taking the derivative with respect to the productivity gap and simplifying yields the following condition for *convergence*

$$\frac{\left(\frac{A_0^{-i}}{A_0^i}\right)^{\delta_K}\left(1+\left(\frac{A_0^{-i}}{A_0^i}\right)^{2\delta_L}\right)+\left(\frac{A_0^{-i}}{A_0^i}\right)^{2\delta_L}\frac{K^i}{K^{-i}}+\left(\frac{A_0^{-i}}{A_0^i}\right)^{2\delta_K}\frac{K^{-i}}{K^i}}{\left(\frac{A_0^{-i}}{A_0^i}\right)^{\delta_K}\left(1+\left(\frac{A_0^{-i}}{A_0^i}\right)^{2\delta_L}\right)\quad+\quad 2\left(\frac{A_0^{-i}}{A_0^i}\right)^{\delta_L+\delta_K}}<\frac{\delta_K}{\delta_L}. \tag{2.41}$$

In the modern economy with A^i being labor productivity so that $\frac{\delta_K}{\delta_L}=\frac{\gamma}{\alpha}\frac{1-\alpha}{1-\gamma}$, condition (2.41) is more likely to hold if $\alpha<\gamma$ (proposition 2.4 below will formalize this). So, world-wide convergence

[3]The parameter choices in figures 2.1 and 2.2 are $\gamma = .65$, $\theta = .5$. In addition, $\frac{K^{-i}}{K^i} = .9$ while $\frac{L^{-i}}{L^i} = 1$ so that region i tends to specialize in agriculture. In figure 2.1, $\alpha = \frac{2}{3}\gamma \approx .43$.

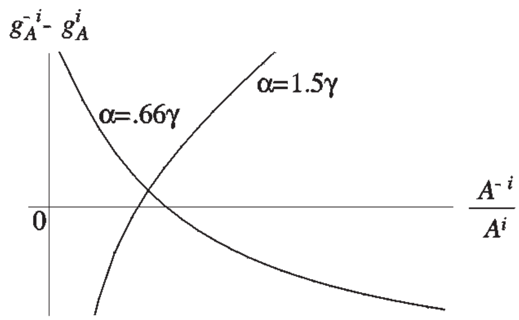


Figure 2.2: Divergence and convergence in the modern economy

in productivity growth is likely to occur if monopoly power in the modern sector is relatively strong or agriculture makes relatively little use of the key factor to growth, or both.

The reason is that monopolistic competition drives a wedge between factor remuneration and factor productivity. The higher monopoly power, the less modern goods $Z^i = N^i z^i = \alpha A^i \lambda^i L^i$ are produced in equilibrium since Z^i is falling in α. Thus, labor is a cheap factor when monopoly power is strong, because the modern sector is small, employs little labor, and the constant-returns-to-scale sector in the background (agriculture) has to employ a lot of labor in general equilibrium. This drives wages down. Simultaneously, monopolistic competition also weakens the specialization force stemming from labor endowments and strengthens the specialization force stemming from capital. The stronger monopoly power gets, that is the further α drops, the less important is the productivity gap in $\Lambda^i = 1 + \left(A^{-i}/A^i\right)^{\alpha}$, and the productivity gap has more impact on $\Lambda^i = 1 + \left(A^{-i}/A^i\right)^{\gamma \frac{1-\alpha}{1-\gamma}} \left(K^{-i}/K^i\right)$. A widening of the productivity gap A^{-i}/A^i strengthens the forces that make region i specialize in the modern sector because agricultural production becomes more attractive in the other region as Γ^i rises. When the productivity gap opens, it has a reverting effect because factor remuneration of skilled workers makes modern production less desirable in a region with a productivity advantage beyond the steady-state level.

In figure 2.1, α is chosen to be relatively small relative to γ ($\alpha = \frac{2}{3}\gamma$). However, for relatively large α ($\alpha = \frac{3}{2}\gamma$), divergence can also occur for modern economies. This is depicted in figure 2.2. Proposition 2.4 states these findings in more general terms.

Proposition 2.4 *After trade liberalization, the four types of economies in table 2.1 (p.25) obey the following dynamics.*

1. *In any 'classical economy', divergence in productivity levels and growth occurs and the region whose specialization forces initially favor agriculture, $\frac{\Lambda^i(t_0)}{\Gamma^i(t_0)} > 1$, completely specializes in agriculture after finite time.*

2. *If there is a steady-state level for which the technology gap between regions i and $-i$ remains constant, it is unique and lies at:*

$$\frac{A_0^{-i}}{A_0^i} = \left(\frac{K^{-i}}{K^i}\right)^{\frac{1}{\delta_L - \delta_K}}. \tag{2.42}$$

   The coefficients δ_K and δ_L are the elasticities of the factor price ratios with respect to the productivity gap, as given in table 2.1 (p. 25).

3. *Local convergence to the steady-state technology gap (2.42) occurs if $\delta_L < \delta_K$. This condition is satisfied in 'modern economies' if A^i is labor productivity and $\alpha < \gamma$. It is always satisfied in 'modern economies' if A^i is total factor productivity.*

Proof. For ease of notation, denote the productivity gap by $a_0 \equiv A_0^{-i}/A_0^i$, the ratio of capital stocks by $k \equiv K^{-i}/K^i$ and the ratio of labor endowments by $l \equiv \frac{L^{-i}}{L^i}$. For a 'classical economy' first consider the case of A^i being labor productivity ($\delta_L = 1$, $\delta_K = 0$). Then condition (2.41) implies that the change in the productivity gap, $g_A^{-i} - g_A^i$, is an increasing function of the productivity gap $a \equiv A^{-i}/A^i$ in a neighborhood of a_0 iff ${a_0}^2 \left[(1+k^{-1}) + (1+k)\right] / \left[1 + 2a_0(1+a_0)\right] \geq 0$. This is always satisfied. Thus, convergence cannot occur in this case. Similarly for the case of a 'classical economy' where A^i is total factor productivity ($\delta_L = \delta_K = 1$), condition (2.41) implies that $g_A^{-i} - g_A^i$ is an increasing function in the productivity gap a at a_0 iff $\left[a_0(1+{a_0}^2) + {a_0}^2(k^{-1}+k)\right] / \left[a_0(1+{a_0}^2) + 2{a_0}^2\right] \geq 1$, i.e. iff $k^{-1} + k \geq 2$. This is always satisfied and convergence cannot occur in this case either. In 'classical economies', a widening productivity gap causes the ratio of specialization forces to increase over time in the region that specializes in agriculture. Taking the partial derivative, one finds $\partial\left(\Lambda^i/\Gamma^i\right)/\partial\left(A^{-i}/A^i\right)\geq 0$ iff $a^{\delta_L-\delta_K}l\left(1+ka^{\delta_K}\right)/k\left(1+la^{\delta_L}\right)\geq\delta_K/\delta_L$. Since $\delta_K/\delta_L=0$ for A^i being labor productivity, this condition is trivially satisfied, and it reduces to $\frac{l}{k}\geq 1$ for A^i being total factor productivity. So, statement 1 in proposition 2.4 holds.

In order to derive the behavior of 'modern economies', observe that the steady-state technology gap is unique if it exists. That is, for $g_A^{-i}=g_A^i$ to hold, (2.40) immediately implies that $k_0=a_0^{\delta_L-\delta_K}$. This proves statement 2 in proposition 2.4.

Moreover, there is a unique zero-intercept of the function $g_A^{-i}-g_A^i$, if it exists. Thus, convergence occurs in a neighborhood of a_0 if convergence condition (2.41) is satisfied at a_0. Using the steady state relationship $k_0=a_0^{\delta_L-\delta_K}$ in condition (2.41) shows that the left-hand side of (2.41) must equal one so that convergence occurs iff $\delta_K/\delta_L>1$. For a 'modern economy', $\delta_L=\alpha$. Using $\delta_K=(\gamma-\alpha\gamma)/(1-\gamma)$ for A^i being labor productivity and $\delta_K=(1-\alpha\gamma)/(1-\gamma)$ for A^i being total factor productivity establishes the third statement in the proposition. ∎

The first statement follows immediately from the fact that the change of the technology gap $g_A^{-i}-g_A^i$ must be an increasing function of the technology gap itself for any 'classical economy'. In general, a steady-state does not need to exist. In addition, the change in the technology gap $g_A^{-i}-g_A^i$ can be a non-monotonic function of A^{-i}/A^i. Figures 2.1b and 2.1d exhibit examples of both. If a steady-state exists, however, then it can only be stable in a 'modern economy'. Using the steady-state value of the technology gap (2.42) in convergence condition (2.41) yields a value of one for the left-hand side of (2.41). Therefore, a steady-state is locally stable iff $\delta_K/\delta_L>1$ or iff $\delta_L-\delta_K<0$. For a modern economy with A^i being labor productivity $\delta_L-\delta_K=\frac{1-\gamma}{\gamma-\alpha}$ so that local convergence occurs for $\alpha<\gamma$. Most surprisingly, a modern economy with A^i being total factor productivity *always* has a stable steady-state to which it converges locally because $\delta_L-\delta_K=\frac{1-\alpha}{1-\gamma}$ is positive. Yet, local convergence is not to be confused with global convergence. As figure 2.1d shows, convergence would not occur in this case if the initial technology gap were severe and A^{-i}/A^i very small.

Policy advice—such as in Rodrik (1999), for instance—stresses the importance of *making openness work* through sound domestic development policies so that an economy can successfully participate in the global marketplace. The results of this section lend support to this view. Countries that are able to participate in intraindustry trade of advanced goods are likely to converge in productivity growth to their trading partners. Once they have successfully prepared themselves for the participation in intraindustry trade, these countries need not fear a negative impact of trade liberalization on their domestic growth. After trade liberalization, dynamic externalities will not widen but reduce the technology gap between South and North to a steady-state level. However, trade will not go further than that. The steady-state level of the technology gap itself can only be reduced through own knowledge creation or knowledge transfers.

2.3.2 P. Romer's (1990) economy

Even though close-form solutions are hard to obtain in more elaborate growth models, it is conceivable that similar forces as in the present model prevail and close the technology gap to a steady-state level. P. Romer's (1990) economy, for instance, shares key features with the economy of this chapter. It has been suggested that free trade between dissimilar regions in the Romer (1990) economy would result in divergence of growth rates. Rivera-Batiz and Romer (1991a) are careful to recommend free trade only for similar regions. It is likely, however, that divergence need not result in their models because of the same reasons for which convergence occurs in the present model.

Compared to the present model, Romer's (1990) economy could be viewed as a one-sector version in which the modern sector consists of three subsectors. Adopting notation used in this chapter, the final production of the modern good in Romer' model takes the form $Z^i = (K_Z)^{1-\zeta} \int_0^{A^i(t)} (z_n^i)^\zeta \, \mathrm{d}n$ where K_Z is some sector-specific factor and z_n^i one variety of an intermediate capital good (for simplicity and without affecting the argument, a third factor in final production was suppressed). The intermediate capital goods are supplied by a continuum of monopolistic competitors, each one selling one variety. In Romer's model, the capital to produce intermediate inputs is supplied through a savings decision and thus given at each instant. For the purpose of this argument it can also be considered a given sector-specific input. Finally, designs for the varieties of capital goods are produced at a rate $\dot{A}^i(t) = B \cdot \lambda^i L^i A^i(t)$, where $\lambda^i L^i$ is the number of workers employed in R&D. Designs are then sold to the intermediate producers at a price so that all rents are shifted to the R&D sector. The modern sector therefore 'suffers' from a monopolistic distortion so that labor demand in R&D is reduced in a similar way to the model in this chapter. This, in turn, distorts factor prices and thus the specialization forces after trade—the first key element that the two models have in common.

A second common feature results once the Romer (1990) model is closed. One can either make the factor K_Z not subsector-specific but also turn it into labor, or one can add a second sector such as agriculture that competes for labor with the R&D sector. To keep the similarities close, follow the second path, set $\zeta = 0$, and add agriculture with a production function $X^i = \left((1-\lambda^i)L^i\right)^\gamma \left(\int_0^{A^i(t)} (z_n^i)\,\mathrm{d}n\right)^{1-\gamma}$. In equilibrium, each variety of capital goods is employed in the same proportion and the two production functions of the Romer model become $X^i = (A^i)^{1-\gamma}\left((1-\lambda^i)L^i\right)^\gamma (\overline{z}^i)^{1-\gamma}$ and $Z^i = A^i(t)\overline{z}^i$. Comparing these production functions to (2.1) and (2.2) shows that the production structures of the two models are closely related while the fact that $\overline{z}^i$ now also enters in agriculture adds an additional source of distortion to the Romer model. Therefore, convergence is likely to occur for large γ just as in the model of the present chapter.

2.4 Agreeing to Diverge: A Dynamic Welfare Analysis

Following P. Romer (1990), Young (1991), Rivera-Batiz and Romer (1991b), Aghion and Howitt (1998, Ch. 10) or Xie (1999), much attention has been paid to a possibly harmful effect of international trade on growth when regions that widely differ in factor endowments or initial productivity start to trade. Some of these approaches, such as Young (1991) or Xie (1999), consider a partial equilibrium, and may therefore miss forces that result in growth convergence as in section 2.3. Moreover, an explicit comparison of welfare losses from slower growth to welfare gains from increased trade seems necessary for normative conclusions. This section provides a dynamic welfare analysis. For this comparison, learning by doing is assumed not to be internalized at all. In addition, only welfare gains that stem from concavity of utility are considered here—neglecting any additional welfare effects from the availability of more varieties. By overstressing dynamic losses and understating repeated static gains in this manner, the model shows that traditional arguments for free trade still carry strong weight, even under conditions of endogenous growth theory.

2.4.1 Repeated static gains

To derive a concise welfare measure, consider output and relative prices first. Lemma 2.1 assembles their levels before and after trade liberalization.

Lemma 2.1 *For all economies in table 2.1 (p. 25), after trade liberalization and* incomplete *specialization output in region i becomes*

$$Y^i_{trade}(t) \quad = \quad \frac{X^i_{trade}(t)}{1-\theta}\left(1+[\theta+\gamma(1-\theta)]\,\frac{\Gamma^i(t)-\Lambda^i(t)}{\Lambda^i(t)}\right) \tag{2.43}$$

in terms of agricultural production, whereas it was

$$Y^i_{aut}(t) = \frac{X^i_{aut}(t)}{1-\theta} \tag{2.44}$$

in autarky. Similarly, the world-wide price index P for modern goods (2.5) becomes

$$P_{trade}(t) \quad = \quad V\,\frac{\Gamma^i}{[A^i(t)^\alpha L^i + A^{-i}(t)^\alpha L^{-i}]^{\frac{1}{\alpha}}}\,\frac{X^i_{trade}(t)}{1-\theta}, \tag{2.45}$$

after trade, whereas it was

$$P_{aut}(t) = V\,\frac{1}{[A^i(t)^\alpha L^i]^{\frac{1}{\alpha}}}\,\frac{X^i_{aut}(t)}{1-\theta} \tag{2.46}$$

in autarky. Here, $V \equiv \frac{1}{\alpha}(L_0)^{\frac{1-\alpha}{\alpha}}[\theta+\gamma(1-\theta)]^{\frac{1}{\alpha}}/(1-\alpha)^{\frac{1-\alpha}{\alpha}}\theta^{\frac{1-\alpha}{\alpha}}$.

Proof. Solving the equation system underlying proposition 2.2 yields

$$Y^i_{trade} = X^i_{trade}\left[(1-\gamma)+\frac{\theta+\gamma(1-\theta)}{1-\theta}\frac{1+\frac{X^{-i}_{trade}}{X^i_{trade}}}{\Lambda^i}\right],$$

which implies (2.43) in the text, for specialization forces as defined in (2.31). The price of modern goods from region j is

$$p^j = \frac{\theta+\gamma(1-\theta)}{\alpha A^j}\frac{X^j}{L^j}\frac{\Gamma^j}{\Lambda^j}.$$

Using the latter relationship along with the equilibrium number of firms in the two regions (2.33) and plugging both into the definition of the price index (2.5) yields (2.45) after a round of simplifications. Similar steps for the autarky equilibrium yield (2.44) and (2.46). ■

Agricultural output after trade liberalization X^i_{trade} is determined by the labor allocation after trade $1-\lambda^i_{trade}$ (proposition 2.2), whereas X^i_{aut} was determined by $1-\lambda^i_{aut}$ (proposition 2.1). An open economy's terms of trade determine its access to wealth and are an important aspect of its welfare. The welfare gains or losses from free trade can be inferred from the representative agent's utility levels. Indirect utility is given by (2.10).

Proposition 2.5 *For all economies in table 2.1 (p. 25), utility attains a level of*

$$\begin{aligned} u^i_{trade}(t) \quad = \quad & T\left(\frac{X^i_{trade}(t)}{1-\theta}\right)^{1-\theta}\cdot\left(1+[\theta+\gamma(1-\theta)]\frac{\Gamma^i(t)-\Lambda^i(t)}{\Lambda^i(t)}\right) \\ & \cdot V^{-\theta}\left(\frac{1}{[A^i(t)^\alpha L^i+A^{-i}(t)^\alpha L^{-i}]^{\frac{1}{\alpha}}}\right)^{-\theta}\left(\Gamma^i(t)\right)^{-\theta} \end{aligned} \tag{2.47}$$

after trade liberalization, while it was

$$u^i_{aut}(t) = T\left(\frac{X^i_{aut}(t)}{1-\theta}\right)^{1-\theta}\cdot 1\cdot V^{-\theta}\left(\frac{1}{[A^i(t)^\alpha L^i]^{\frac{1}{\alpha}}}\right)^{-\theta} \tag{2.48}$$

in autarky. Thus, the ratio of post- and pre-trade utility becomes

$$\begin{aligned} \frac{u^i_{trade}}{u^i_{aut}}(t) \quad = \quad & \left((1-\gamma)(1-\theta)\left(\frac{\Lambda^i(t)}{\Gamma^i(t)}\right)^{\theta+\gamma(1-\theta)}\right. \\ & \left.+[\theta+\gamma(1-\theta)]\left(\frac{\Gamma^i(t)}{\Lambda^i(t)}\right)^{(1-\gamma)(1-\theta)}\right)\cdot\left(\Lambda^i(t)\right)^{\frac{1-\alpha}{\alpha}\theta}. \end{aligned} \tag{2.49}$$

Proof. Using the results of lemma 2.1 in indirect utility $u^i = TY^iP^{-\theta}$ from (2.10) yields (2.47) and (2.48). Dividing (2.47) by (2.48) and simplifying yields (2.49). ■

The three terms in (2.47) have intuitive interpretations. Neglecting the constant T, the first factor $\left(X^i_{trade}/1-\theta\right)^{1-\theta}$ captures the *reallocation effect* of trade liberalization. Depending on the specialization forces that shift the labor allocation under free trade, this term is larger or smaller than the corresponding term under autarky. It indicates whether agricultural output increases or falls after trade liberalization. The *reallocation effect* works through both output and prices as can be seen from (2.43) and (2.45). Hence, the power of $1-\theta$.

The second term in (2.47) only appears in utility after trade, but not in utility before trade. One could call it, somewhat euphemistically, the '*output effect*'. In fact, this is no real effect because the true *reallocation effect* was captured entirely by the previous term. The nominal '*output effect*' arises because the agricultural good is the *numéraire*. Were the modern good the *numéraire*, the effect would work differently.

The third term in (2.47) exclusively captures the *price effect* of trade liberalization. Its moves incorporate the terms-of-trade effect: A slowly growing region can share in the growth of the other region through an improvement of its terms of trade. Depending on whether the foreign region can produce more productively than the home region, and depending on how strongly the specialization force from capital endowments Γ^i shifts employment, the average price of modern goods will fall or rise from the view point of region i. Only the product of all three effects, the *reallocation*, '*output*' and *price effect*, correctly accounts for the increase in welfare. The *price effect* will offset nominal '*output effects*' so that the true welfare gains from trade are captured. The analysis highlights that, in an open economy, welfare gains of the representative agent cannot be inferred directly from changes in output.

An economy gains from trade liberalization in static terms iff u^i_{trade}/u^i_{aut} exceeds unity in (2.49). In 'classical economies', $\alpha = 1$ so that the factor $\Lambda^i(t)^{\frac{1-\alpha}{\alpha}\theta}$ in expression (2.49) drops out.

The term captures the utility that consumers derive from varieties in 'modern economies'. Since these gains stem from the particular form of utility that was imposed and go beyond gains from concavity, these extra-gains will be disregarded in the following arguments.

The first factor in (2.49) expresses conventional static gains from trade. For $\Lambda^i(t) \neq \Gamma^i(t)$, this term strictly exceeds unity for it takes the form $a \cdot x^{1-a} + (1-a) \cdot x^{-a}$, and $a \in (0,1)$, $x \in (0,\infty)$. Since both specialization forces, $\Lambda^i(t)$ and $\Gamma^i(t)$, always exceed unity, their ratio cannot become negative and $x \in (0,\infty)$. Figure 2.3 plots the utility ratio as a function of the specialization forces. Conventional gains from the first factor in (2.49) are depicted by the thick curve. There are no gains from trade when the two specialization forces are exactly offsetting so that their ratio equals one. As is well known since David Ricardo, there are no static losses from trade. In fact, as the concave curve in figure 2.3 illustrates and differentiation shows, the gains from trade increase more than proportionally when the two regions become less similar.

If the representative agent in country i only considered these static gains, a region would always choose to liberalize trade. The horizontal axis measures zero-gains from trade and represents the utility level in autarky. Any move away from unity on this axis increases welfare beyond the autarky level. But what if there are dynamic losses that offset the static gains? What if the two countries diverge after free trade and the productivity gap A^{-i}/A^i widens?

2.4.2 Dynamic welfare analysis

To analyze the trade-off between repeated static gains from trade and potential dynamic losses, a dynamic criterion beyond (2.49) is needed. When deciding whether to open up to free trade at time t_0, the representative agent faces the choice between receiving the autarky utility forever, $\int_{t_0}^{\infty} e^{-\rho(\tau-t)}\, u^i_{aut}(\tau)\, \mathrm{d}\tau$, or the utility from free trade forever $\int_{t_0}^{\infty} e^{-\rho(\tau-t)}\, u^i_{trade}(\tau)\, \mathrm{d}\tau$. The specialization forces in this model are not reverting or cyclical so that the agent is not concerned about temporary trade liberalization. If it pays to liberalize at some point t_0, then it always pays to liberalize.

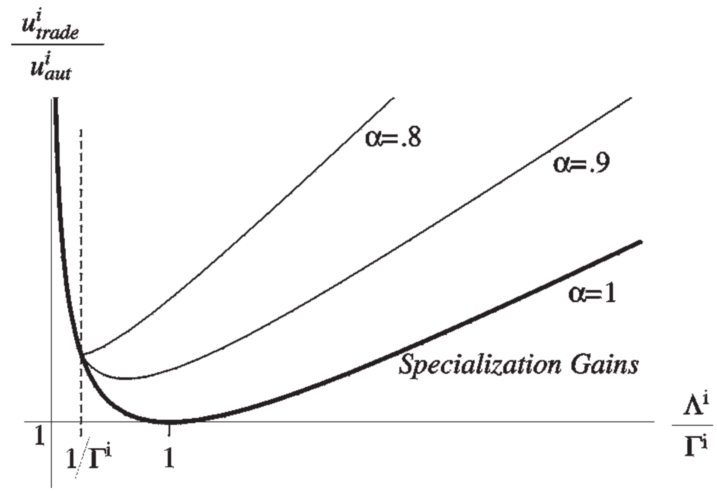


Figure 2.3: Gains from trade

Corollary 2.5.1 *The appropriate welfare criterion for trade liberalization under dynamic considerations is*

$$\left[\rho - (\theta + \gamma(1-\theta))g^i_{A,aut}\right] \int_{t_0}^{\infty} e^{-\int_{t_0}^{\tau}\left[\rho-(\theta+\gamma(1-\theta))g^i_{A,trade}(s)\right]ds}\, \frac{u^i_{trade}}{u^i_{aut}}(\tau)\, d\tau > 1. \tag{2.50}$$

Proof. Expression (2.50) follows from

$$\begin{aligned} & \frac{\int_{t_0}^{\infty} e^{-\rho(\tau-t_0)}\, u^i_{trade}(\tau)\, \mathrm{d}\tau}{\int_{t_0}^{\infty} e^{-\rho(\tau-t_0)}\, u^i_{aut}(\tau)\, \mathrm{d}\tau} > 1 \\ \Leftrightarrow \quad & \frac{\int_{t_0}^{\infty} e^{-\rho(\tau-t_0)}\, \left[A^i_{trade}(\tau)\right]^{\theta+\gamma(1-\theta)}\, \frac{u^i_{trade}}{u^i_{aut}}(\tau)\, \mathrm{d}\tau}{\int_{t_0}^{\infty} e^{-\rho(\tau-t_0)}\, \left[A^i_{aut}(\tau)\right]^{\theta+\gamma(1-\theta)}\, \mathrm{d}\tau} > 1. \end{aligned}$$

The first equivalence holds because labor and capital endowments do not change over time. The equivalence with (2.50) follows from the facts that $A^i(\tau) = A^i(t) \cdot e^{\int_{t_0}^{\tau} g^i_A(s)\mathrm{d}s}$ and $g^i_{A,aut} = const.$ It requires that $\frac{\rho}{\theta+\gamma(1-\theta)} > \int_{t_0}^{\tau} g^i_A(s)\, \mathrm{d}s$, otherwise the integrals do not take a finite value. ■

If (2.50) exceeds unity, the region is better off liberalizing. If it is less than one, the region prefers to remain autark. To gain an understanding of the criterion, consider the extreme case in which region i completely specializes in agriculture after trade and stops growing ($g^i_{A,trade} = 0$). If there were no static gains from trade, criterion (2.50) would become $\rho - g^i_{A,aut} < \rho$ and the representative agent would refuse to liberalize to free trade. There are static gains from trade, however. They occur each period and that forever. Over time, they sum up to large benefits. If the specialization forces did not change over time, a sufficient condition for trade liberalization would be

$$\begin{aligned}\frac{u^i_{trade}}{u^i_{aut}} &= (1-\gamma)(1-\theta)\left(\frac{\Lambda^i}{\Gamma^i}\right)^{\theta+\gamma(1-\theta)} + [\theta+\gamma(1-\theta)]\left(\frac{\Gamma^i}{\Lambda^i}\right)^{(1-\gamma)(1-\theta)} \\ &> \frac{\rho}{\rho - g^i_{A,aut}}. \end{aligned} \tag{2.51}$$

This implicitly defines a lower bound on the specialization force ratio Λ^i/Γ^i beyond which free trade is desirable. If the specialization force ratio is not fixed but grows over time, as is the case for all 'classical economies', the lower bound beyond which trade is desirable will be even smaller because additional future gains from trade make up for some of the losses in growth.

This lower bound on the specialization force ratio, however, varies heavily with the choice of the discount factor and the autarky growth rate. Therefore, a more useful quantity to look at is the ratio $g^i_{A,aut}/\rho$. For (2.51) to be satisfied, this ratio must not exceed $1 - u^i_{aut}/u^i_{trade}$. To evaluate u^i_{aut}/u^i_{trade}, one can choose the point of complete specialization in agriculture as reference value: $\Lambda^i_0/\Gamma^i_0 = 1 + \theta/[\theta+\gamma(1-\theta)]$. At that point, region i stops growing and incurs the worst dynamic loss possible. Table 2.2 reports values of $1 - u^i_{aut}/u^i_{trade}$ for different parameter choices. As long as the ratio $g^i_{A,aut}/\rho$ is less than or equal to the values reported, region i prefers free trade. Suppose region i would grow at a rate of 5% in autarky. Then, the representative agent has to be at least patient enough so that $\rho \geq .07$ for $\theta = .7$ and $\gamma = .1$, or $\rho \geq .22$ for $\theta = .5$ and $\gamma = .3$. For $\theta = .3$ and $\gamma = .7$, however, the agent would have to be extremely patient with $\rho \geq 2.45$.

The estimates are conservative. A region chooses free trade if the condition in table 2.2 is satisfied but not only in that case. First, to keep calculations simple, the more advanced foreign region $-i$ was assumed not to grow after trade liberalization. The model, however, predicts increased growth in the more advanced region after trade liberalization and the less advanced region can share in this growth through improving terms of trade. Second, the less advanced region was assumed to

Table 2.2: Lower Bounds on Discounted Growth for Trade Liberalization to be Desirable

$(1 - \frac{u_{aut}}{u_{trade}} \Big\vert_{\frac{\Lambda}{\Gamma}=1+\frac{\theta}{\gamma(1-\theta)}})$	$\theta = .3$	$\theta = .5$	$\theta = .7$
$\gamma = .1$	.228	.464	.676
$\gamma = .3$	.091	.229	.401
$\gamma = .5$	.045	.123	.237
$\gamma = .7$	.020	.059	.122

Criterion: Free trade iff $\frac{g_{A,aut}}{\rho} \leq 1 - \frac{u_{aut}}{u_{trade}} \Big\vert_{\frac{\Lambda}{\Gamma}=1+\frac{\theta}{\gamma(1-\theta)}})$

stop growing immediately. However, there is a transition period of slowing but non-zero growth. On both accounts, the LDC would benefit. Finally, an average growth rate of 5% *forever* is high even for successful developing regions (Sachs and Warner 1995).

So, for a reasonably broad range of parameter values, regions strictly prefer trade over autarky even if they subsequently grow more slowly. The calculations have an immediate implication especially for a poor economy with relatively low productivity levels and autarky growth prospects. Call such a place Antarctica, say. If potential trading partners raise their productivity at a relatively fast rate, the specialization forces that would prevail if Antarctica opened up to free trade get stronger and stronger. They surpass the boundary level at some point. Thus, a country with lastingly low growth prospects will always agree to trade liberalization after some point in time, and prefer trade and divergence over autarky. In this sense, the less developed region *agrees to diverge* while the faster growing region welcomes divergence anyway. Even fast growing less developed regions may prefer free trade over isolation if the factor endowments between the regions differ so strongly that the gains from trade outweigh dynamic losses from divergence. These countries will *agree to diverge* as well.

The cases considered are worst-case scenarios for LDCs. The only source of growth is a not even partly internalized dynamic externality. Even under such conditions, the dynamic losses from trade may be small compared to repeated static gains. In practice, own R&D efforts and knowledge transfers help nurture innovative and growth promoting sectors so that the worst case is unlikely to apply.

2.5 Conclusion

What do new trade and new growth theory imply for welfare and growth convergence after trade liberalization? Four types of economies were considered in this chapter. They exhibit two types of productivity change—labor and total factor productivity growth—and two types of competition in the sector that determines growth—perfect competition under constant returns to scale and monopolistic competition under increasing returns to scale. Both perfect competition and monopolistic competition leave no rents to be shifted across regions. So, industry location does not matter in this static sense. It matters heavily, however, for the dynamic externality from learning by doing. When labor is immobile across regions, the more modern firms a region hosts the more it benefits from this growth externality.

The main insights of this chapter are twofold. First, and in contrast to previous arguments in new growth theory, if the innovative sector is characterized by monopolistic competition, then free trade can result in international growth convergence. The reason is that monopolistic competition can reverse the forces of specialization that prevail in classic trade theory. Countries that manage to participate in intraindustry trade for advanced goods after trade liberalization are likely to converge to the growth rates of richer countries. Specialization in low-growth sectors will not make these countries fall behind.

Second, gains from international trade and specialization are static gains, repeatedly realized in every instant, and sum up to large benefits over time. These benefits can outweigh dynamic losses from slower growth after trade liberalization, and countries may choose free trade over isolation even if trade causes divergence. In addition, the slower growing regions can share in the wealth creation of the faster growing regions through improving terms of trade. As a consequence, trade and divergence can be better than isolation for substantial ranges of parameters.

Even though these results speak strongly for trade liberalization, it should not be seen as an unconditionally desirable policy. In the light of the convergence result, a country may also choose to pursue temporary trade restrictions, prepare its domestic industries for their successful participation in intraindustry trade, open up to free trade as soon as the modern sector is able to compete successfully with foreign firms, and then benefit from convergence under international trade. In this sense, developing countries can engage world markets on their own terms.

Chapter 3

Trade and productivity change: A study of Brazilian manufacturers, 1986-1998

The preceding chapter 2 revisited trade theory in its relationship to growth. Theoretical contributions addressing this concern abound. However, the precise channels through which trade affects growth seem to remain little understood. Does an open economy grow faster than a closed economy and, if so, why? Over the past two decades, many developing countries and transition economies have liberalized international trade. In the early nineties, Brazil reduced tariffs and eliminated import barriers. Among other motives, Brazil aimed to promote productivity gains by exposing firms to fiercer international competition and facilitating access to global markets.

Different branches of empirical research investigate how more open economies fare. One line of empirical studies approaches the question through cross-country comparisons. For the most part, findings indicate that open countries tend to grow faster (e.g. Ben-David 1993, Sachs and Warner 1995). However, these studies have spawned criticism for the possible arbitrariness of indices of openness and endogeneity problems (Slaughter 1997, Harrison and Hanson 1999, Rodríguez and Rodrik 2000). Sector-level studies circumvent these problems but cannot unmask underlying microeconomic processes either (e.g. Keller 2000, Kim 2000). Using longitudinal data, a microeconomic branch of empirical research traces the effects of trade exposure and foreign direct investment on firms or plants in select countries. Caves (1998), Bartelsman and Doms (2000), and Tybout (2001) review the microeconometric literature on industry turnover and productivity change. The study of this chapter belongs in this latter branch, too.

For the first time, the present study employs an unbalanced panel of 9,500 medium-sized to large Brazilian manufacturers between 1986 and 1998. Beyond commonly available variables, firms report foreign equipment acquisitions and the use of foreign intermediate inputs. Special variables trace a firm's economic destiny—its state of operation, and its suspension or extinction. These groups of variables permit refinements in the estimation technique that were not feasible with previous data.

Brazil strongly reduced import barriers during the early nineties. Exports, however, remained subject to largely unaltered taxes and tariffs between 1986 and 1998 (Veiga 1998). As a welcome consequence, the impact of trade reform on the import side can be isolated from other effects of trade. Foreign direct investment, a further key aspect of an economy's openness, rose strongly in Brazil over the same period (Bonelli 1999) and will be controlled for. This chapter asks: *How did Brazil's removal of import barriers affect productivity among its medium-sized to large manufacturers?*

Trade reform may affect productivity through three primary *channels*.

1. *Foreign Input Push*: High-quality equipment and intermediate goods allow firms to adopt new production methods. Their use can raise efficiency.

2. *Competitive Push*: The removal of import barriers increases competition on the product market side. This can allow firms to remove agency problems and induce them to innovate processes.

These two effects tend to shift a *firm's* productivity. In addition, a separate group of trade effects on productivity can only be observed at the level of *sectors or industries*. The present analysis focuses on

3. *Competitive Elimination*: Increased foreign competition makes the least efficient firms shut-down and enables the surviving, competitive firms to increase market share. This turnover raises average productivity.

The present dataset allows for the separation of the foreign input push (1). Feenstra, Markusen and Zeile (1992, Korean business groups) and Fernandes (2001, Colombian manufacturers) also trace effects of inputs on productivity at the micro-level. Their studies suggest that productivity is positively related to the use of high-quality inputs. The present study shows, however, that this effect is relatively small compared to the other two channels. Microeconometric studies on the competitive push (2) and competitive elimination (3) include Cox and Harris (1985), Levinsohn (1993), Roberts and Tybout, eds (1996) and Pavcnik (2000). For Brazil, Cavalcanti Ferreira and Rossi Júnior (1999) find a positive impact of trade reform on productivity in sector-level data and Hay (2001) in a sample of 320 large manufacturers. In general, these studies show that higher efficiency and faster turnover follow trade liberalization. However, as Tybout (2001, p. 16) concludes in a recent literature review, "it is difficult to find studies that convincingly link these processes to the trade regime." The present chapter aims to establish causal relationships between productivity and trade exposure, and sets out to evaluate the relative importance of the above three channels.

Like previous studies, the chapter employs Olley and Pakes' (1996) estimation algorithm to obtain consistent productivity estimates. A new structural model motivates the procedure. The model resolves a previous shortcoming, which demanded that observations with zero (negative) investment be excluded. This would have caused the loss of up to a third of all observations. Productivity estimates are corrected for the endogeneity of price in sales figures (Klette and Griliches 1996) and for changing managerial efficiency under increasing competition. Firm-level productivity estimates and turnover are then related causally to the trade regime.

Results for the foreign input push (1) indicate that, on average, the efficiency of foreign equipment and intermediate inputs is higher than the efficiency of domestic inputs. To measure their effect, foreign inputs are included in the production function and distinguished regarding their role as capital goods or intermediate inputs. However, their overall efficiency contribution is minor. The adoption of new technologies can reduce productivity initially. Firms need to put high-quality inputs to adequate use in order to achieve productivity gains. In several sectors, Brazilian firms do not appear to succeed with necessary rearrangements in the short term.

Evidence regarding the competitive push (2) indicates that firms respond strongly to increased competitive pressure and raise their efficiency. To draw this causal conclusion, the analysis employs instrumental variables (IV) and controls for the endogeneity of foreign market penetration and trade policy. Third, firm turnover and the exit of the least productive firms contributes positively to productivity change in the aggregate. In an effort to evaluate this competitive elimination (3) causally, probabilities of Markov transitions between states of operation are estimated as a function of the trade regime. The exit probability increases strongly with foreign competition.

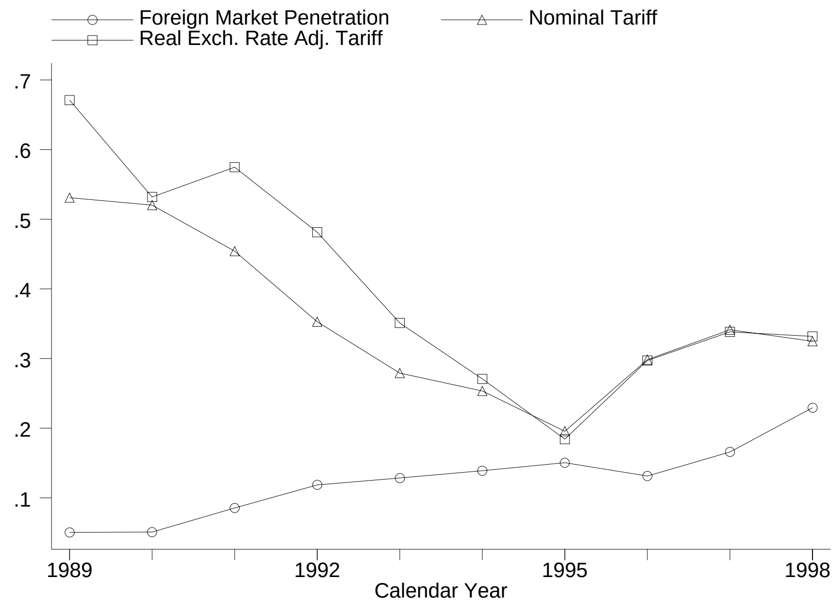
To understand the relative importance of the three channels *vis à vis* each other, counterfactuals are evaluated in simulations. The counterfactuals ask how much less productivity change would have occurred through each channel had Brazil not reduced tariffs. These simulations show that the competitive push (2) is a salient source of immediate productivity change, while competitive elimination (3) exerts its impact slowly.

The remainder of the chapter is divided in six sections. Section 3.1 gives a short overview of the Brazilian trade liberalization programme. The data are briefly described in section 3.2, while details are relegated to appendix A. Section 3.3 presents the behavioral framework. It bases Olley and Pakes' (1996) estimation procedure on different theoretical grounds and adapts the technique to the present context. Section 3.4 documents how consistent firm-specific productivity measures are obtained. Using the resulting firm-level productivity estimates, section 3.5 evaluates how Brazil's trade policy affected productivity and distinguishes the three above-mentioned channels. Section 3.6 is the conclusion.

3.1 Brazilian Trade Policy

For decades, policies of import substitution and industry protection were part of Brazil's broader development strategy. Until the early nineties, high tariffs, exchange rate controls and interventions, and especially prohibitive non-tariff barriers were intended to reduce the competitive pressure on domestic industries (Bonelli, Veiga and Brito 1997). From the mid seventies until the late eighties, for example, potential importers to Brazil had to undergo a rigorous examination as to whether their commodities were similar to domestic products. If so, their imports were banned. As a result, the Brazilian domestic market remained essentially closed for a broad range of foreign equipment, including computers.

In 1988, the federal government initiated a process of trade liberalization that reduced both the level and the cross-industry dispersion of tariffs. Redundant tariffs were eliminated from Brazil's tariff act by 1990, but the binding non-tariff barriers remained largely untouched. Only the Collor de Melo administration in 1990 was able to break with earlier Brazilian policies of import substitution and industrial targeting (Bonelli et al. 1997). The government presented a detailed schedule for tariff reductions to be completed by 1994 and announced the elimination of non-tariff barriers (Horta, Piani and Kume 1991). Tariffs on equipment not produced in Brazil, for instance, were immediately reduced to zero and non-tariff barriers were eliminated. At the other extreme, tariffs for information technology remained at 40 percent in order to protect Brazil's fledgling computer industry. In fact, the liberalization programme was concluded in less than three years by July 1993. This speed and the far reaching removal of non-tariff barriers shocked the domestic manufacturing sector considerably. However, when president Cardoso took office in 1995, some of the earlier liberalization efforts were reversed.



Source: Own calculations, tariffs weighted by imports. Sector: *Bens de capital, equipamentos de transporte* (definition according to Mesquita Moreira and Correa 1997).
Data: Ramos and Zonenschain (2000a) for market penetration; Kume, Piani and Souza (2000) for nominal tariffs; Brazilian central bank and US census bureau for real exchange rate.

Figure 3.1: Tariffs and foreign market penetration. Car makers

The transport equipment sector is a striking example. Tariff levels and market penetration are depicted in figure 3.1. Brazil's elevated real exchange rate added to protection until 1994. To show this, a tariff series weighted by the real exchange rate is depicted alongside. Foreign market penetration tends to mirror the moves of Brazil's effective tariffs fairly closely. Machinery, toy and car makers were among the most successful lobbyists to argue for an increase in tariffs in 1995. Other sectors faced less of a reversal.

Cavalcanti Ferreira and Rossi Júnior (1999) argue that, on average, the effective rate of protection was about 80 percent of the import price in 1985. They control for both tariffs and non-tariff barriers. According to their measure, effective protection fell to 21 percent by 1997—a quarter of the level twelve years earlier. Brazil took hardly any steps to stimulate exporting (Veiga 1998). Select sectors benefit from largely unaltered export incentives. However, peculiar double-taxation schemes could and still can make exporting more expensive for firms in most sectors than selling to the domestic market.

3.2 Data

The Brazilian census bureau IBGE surveys manufacturing firms and plants annually in its *Pesquisa Industrial Anual* (*PIA*). Plant data in *PIA* are considerably less comprehensive, so the present study only draws on firm data. In addition, firm data are more likely to capture unobserved characteristics and inputs such as managerial ability than are plant data. The current sample is not strictly representative for the manufacturing sector as a whole. Yet, to trace the effects of trade liberalization on productivity, only a random sample is needed that was selected independent of trade exposure. This is satisfied.

A firm qualifies for *PIA* if at least half of its revenues stem from manufacturing activity and if it is formally registered with the Brazilian tax authorities. In 1986, the initial *PIA* sample was constructed from three layers: (1) A non-random sample of the largest Brazilian manufacturers with output corresponding to at least 200 million Brazilian Reais (BRL) in 1995 (around 200 million US dollars in 1995). There were roughly 1,000 of them. (2) A random sample among medium-sized firms whose annual output in 1985 exceeded a value corresponding to BRL 100,000 in 1995 (around USD 100,000 in 1995). About 6,700 firms made it into *PIA* this way. (3) A non-random selection of newly founded firms. *PIA* only included new firms that surpassed an annual average employment level of at least 100 persons. The inclusion process ended in 1993, however. Until then, around 1,800 firms were identified in this manner.

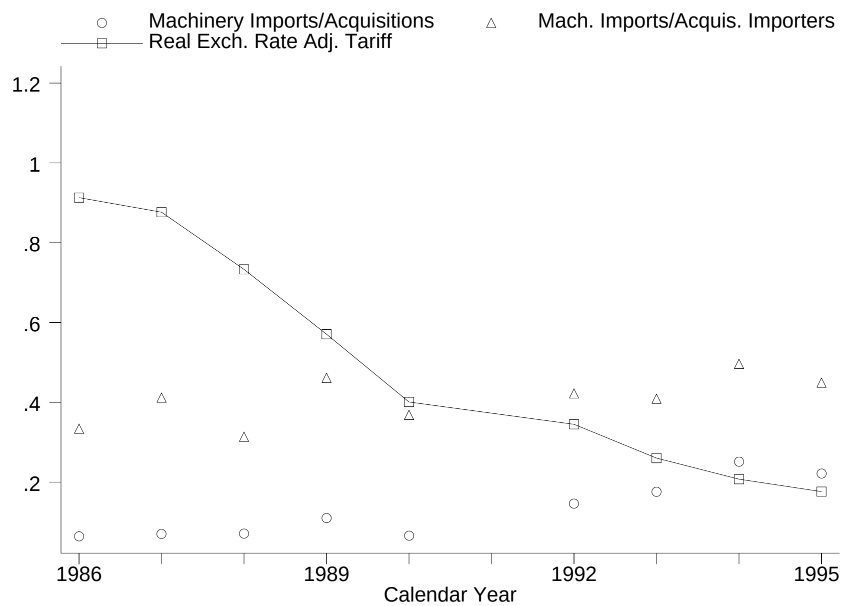
Departing from this initial sample, *PIA* identifies more than 9,500 active firms over the years. A firm that ever enters *PIA* through one of the selection criteria remains in the sample unless it is legally extinct. Moreover, if an existing firm in firm in *PIA* reports the creation of a new firm as a subsidiary or spin-off, this new firm enters *PIA*, too. No sample was taken in 1991 due to a federal austerity program. The sampling method changed in 1996, and no capital stock figures are reported since. Therefore, the dataset of this chapter only embraces firms after 1995 that were present in *PIA* earlier or that were longitudinally related to an earlier firm. Their capital stock is inferred with a perpetual inventory method. Following the change in methods, there is a drop in the sample in 1996. Tests at various stages of the estimation prove it exogenous.

PIA includes precise longitudinal information for every firm. Special variables summarize a firm's state of operation and make sure that observations with missing economic information are not confused with a shutdown or a temporary suspension of production. This is particularly important since it was quite common among Brazilian manufacturers between 1986 and 1998 to *mothball* for extended periods of time. Among the 9,500 firms, more than 1,100 state in at least one year that they suspended production temporarily or for the entire year. The construction of an unbalanced panel from *PIA* is documented in appendix A.

Economic variables in *PIA* include sales figures and changes in final goods stocks, costs of inputs, salaries, employment of blue- and white-collar workers, and several variables related to investment and the capital stock. Most interestingly, firms in *PIA* report their acquisitions of foreign equipment until 1995 and their purchases of foreign intermediate goods since 1996. These foreign acquisitions are reported with their prices at the time of purchase. So, prices properly reflect the goods' relative economic values at which the firms decide to buy. I describe in detail in appendix A how coherent capital stock and foreign equipment series are obtained. Foreign equipment shares are inferred from foreign equipment investments and overall retirements.

Figure 3.2 plots the evolution of foreign equipment acquisitions between 1986 and 1995. Importers of foreign equipment before 1991 continue to invest in foreign equipment at roughly the same rate after 1991. However, the share of foreign equipment in total equipment acquisitions jumps up significantly. This suggests that mostly firms that did not acquire foreign equipment before 1991 do so after trade liberalization.

Sector classifications in *PIA* would allow for the estimation of production functions at a



Source: Own calculations.
Data: *Pesquisa Industrial Annual* for equipment acquisitions. Effective equipment tariffs from Kume et al. (2000) weighed by the national capital formation vector (*IBGE*).

Figure 3.2: Foreign equipment acquisitions. All manufacturing

level that corresponds to three SIC digits (*nível 100*). However, large firms in *PIA* are likely to offer product ranges beyond narrowly defined sector limits. Data at more aggregate levels also provide more variation in the cross section because market penetration and tariff series then become available for two or more subsectors within several sectors. Moreover, switching from the three to the two-digit level increases the number of observations per estimation considerably. So, estimations are carried out at the SIC two-digit level (*nível 50*).

3.3 Behavioral Framework

Three main types of biases affect estimates of production with microdata. First, there is self-selection. Surviving firms exhibit higher productivity than less competitive exiting firms. So, an unbalanced panel of firms needs to be constructed.

Second, firms choose investment and decide whether to stay in business given their productivity. However, productivity is not observable directly. Marschak and Andrews (1944) outline this simultaneity problem early on. It has become known as 'transmission bias'. Such a bias may affect the capital coefficient for two reasons. For one, firms with a large capital stock are more likely to remain in business and tolerate lower productivity levels. This can introduce a negative bias in the capital coefficient as Olley and Pakes (1996) show. On the other hand, if firms can invest both in capital goods and in productivity, they may choose to raise or let decay capital and productivity simultaneously. The model below will clarify that this can introduce a positive bias. Theoretically, either source can dominate. Some sectors in the present data exhibit a positive bias, while others suffer from a negative bias.

Third, error in production function estimates and hence productivity measures arises from an 'omitted price bias' as discussed by Klette and Griliches (1996). This problem occurs when sales figures are used to approximate output. It has also been known since at least Marschak and Andrews (1944).

The present section develops a behavioral framework suitable to address these issues. The model resolves a previous shortcoming of Olley and Pakes (1996) that called for observations with zero (negative) investment to be dropped from the sample. In addition, the model provides a production-side explanation why the bias in the capital coefficient may be positive, a frequent empirical finding. The model motivates Olley and Pakes' estimation procedure differently and incorporates Klette and Griliches' (1996) correction method. Its main feature is that firms can invest in both capital goods and productivity. Finally, the model permits that competition may induce managers to remove inefficiencies—a maintained hypothesis of the present study. Optimality conditions are derived using principal-agent and q-theory, and Olley and Pakes' regression equations will be based on those conditions. Two main testable implications of the proposed model are that the productivity level and the capital stock should be positively correlated and that productivity should be procyclical. Both implications are borne out in the present data (the level regressions in

Table 3.1: Components of q-Theory Model

Variable	Evolution / Timing	Observed	Olley & Pakes
State Variables			
TFP: $(\Omega_{i,t})^{\nu}$	$\Omega_{i,t} = \left[\Omega_{i,t-1}(1-\delta^{\Omega}) + I^{\Omega}_{i,t-1}\right] \tilde{x}_{i,t}$	no	Markovian[a]
Capital $K_{i,t}$	$K_{i,t} = K_{i,t-1}(1-\delta^{K}) + I^{K}_{i,t-1}$	yes	same
Control Variables			
Investment $I^{\Omega}_{i,t-1}$	before $\tilde{x}_{i,t}$ realized (based on $q^{\Omega}_{i,t}$)	no	
Investment $I^{K}_{i,t-1}$	before $\tilde{x}_{i,t}$ realized (based on $q^{K}_{i,t}$)	yes	same
Survival $\chi_{i,t}$	after $\tilde{x}_{i,t}$ realized	yes	same
Labor $L_{i,t}$	after $\tilde{x}_{i,t}$ realized	yes	same
Implications			
Upward bias in capital coefficient explained			
Observations with zero (negative) investment permissible			no

[a] Olley and Pakes (1996) consider an exogenous Markov process of TFP beyond a firm's control. Alternatively, Ericson and Pakes (1995) allow for a binary choice of TFP improvement that affects the Markov process. The current model allows managerial effort to alter the distribution of $\tilde{x}_{i,t}$ as in a standard principal-agent model. Whereas $I^{\Omega}_{i,t}$ is observable to a firm's owner from cash flows, managerial effort is not.

table 3.8 are examples).

3.3.1 Assumptions

Firms can invest in two state variables: capital and total factor productivity (TFP). There are several flow variables. Besides investment, which moves the two state variables, firms employ labor and use intermediate goods. For ease of exposition, consider just one type of capital and labor for now and neglect intermediate inputs. Table 3.1 gives an overview of the main ingredients and the consequences of the model to be derived.

The variable $\Omega_{i,t}$ is the total of a firm's tacit knowledge, organizational skills, and efficiency-relevant arrangements embodied in the production process. All of these factors contribute to a firms' TFP level. They are not transferrable from one firm to another but can be accumulated within a firm. They depreciate unless cultivated with investment $I^{\Omega}_{i,t}$. For simplicity, TFP is assumed to be

$$\mathrm{TFP}_{i,t} = (\Omega_{i,t})^{\nu} \tag{3.1}$$

for some coefficient $\nu > 0$. As opposed to physical capital accumulation, there is a stochastic factor $\tilde{x}_{i,t}$ to the evolution of organizational knowledge:

$$\Omega_{i,t} = \left[\Omega_{i,t-1}(1-\delta^{\Omega}) + I^{\Omega}_{i,t-1}\right] \cdot \tilde{x}_{i,t}. \tag{3.2}$$

The parameter δ^{Ω} expresses the depreciation rate of organizational knowledge. Productivity choice is an imperfect substitute for physical capital because $(\Omega_{i,t})^{\nu}$ will enter the production function separately and a firm cannot anticipate the realization $x_{i,t}$. The stochastic factor $\tilde{x}_{i,t}$ captures a firm's efficiency and is assumed to be uncorrelated with its past realizations and factor inputs—similar to the spirit of Olley and Pakes' model. However, the efforts of a firm's management to improve efficiency and make better use of organizational skills can affect the distribution of $\tilde{x}_{i,t}$ favorably (more on this in subsection 3.3.3).

Consider a market with monopolistic competition. Each firm manufactures one variety of a good. Consumers have income Y_t and preferences as in a standard model for intraindustry trade: $u(Z_1, ...Z_N; C) = (\theta/\alpha)\ln(\sum_{n=1}^{N}(Z_n)^{\alpha}) + (1-\theta)\ln C$. There are N varieties of good Z. Under this utility, price elasticity of demand for a modern good i is approximately $-1/(1-\alpha)$. With a price index $\bar{P}_t \equiv [\sum_{n=1}^{N} P_{n,t}^{-\alpha/(1-\alpha)}]^{-(1-\alpha)/\alpha}$, similar to a census bureau's price index, demand for firm i's good can be stated as

$$D_{i,t} = \frac{\Theta_t}{\bar{P}_t} \cdot \left(\frac{P_{i,t}}{\bar{P}_t}\right)^{-\frac{1}{1-\alpha}}, \tag{3.3}$$

where $\Theta \equiv \theta Y_t$ is the income share that domestic consumers spend on goods Z, including imports. This will be a key relationship for the correction of endogenous price in sales (Klette and Griliches 1996).

To see more clearly how foreign competition affects demand, suppose that domestic varieties of a good sell at about the same price. However, there is a possibly different world market price P_t^f for foreign varieties that compete with firm i's good. Then, domestic demand for a domestic manufacturer i's variety can be approximated by[1]

$$D_{i,t} = \frac{1}{1 + \frac{N_t^{for}}{N_t^{dom}} \left(\frac{P_{i,t}}{\varepsilon_t P_t^f (1+\tau_{i,t})} \right)^{\frac{\alpha}{1-\alpha}}} \frac{\Theta_t}{N_t^{dom} P_{i,t}}, \tag{3.4}$$

where ε_t is the nominal exchange rate, and $\tau_{i,t}$ the nominal tariff in the market of firm i. N_t^{dom} and N_t^{for} denote the number of domestic and foreign varieties, respectively. Their ratio is a measure of foreign market penetration. Demand for a domestic firm's variety increases when there are relatively fewer foreign competitors, or when foreign price is higher, tariffs are higher, or the exchange rate is more favorable—as one would expect.

3.3.2 A firm's price, factor and investment choice

A monopolist in the market for good Z sets price and chooses the variable factors in every period t, given his capital stock and TFP. The production technology for variety i of good Z is assumed to be Cobb-Douglas:

$$Z_{i,t} = (\Omega_{i,t})^{\nu} (K_{i,t})^{1-\beta} (L_{i,t} - L_0)^{\beta}, \tag{3.5}$$

where $(\Omega_{i,t})^{\nu}$ is TFP. $L_{i,t}$ denotes employment, the only variable factor for now. L_0 is the fixed labor input and needed in every period to keep the firm in operation. It gives rise to monopolistic competition in equilibrium.

Consider a firm's intertemporal choice of its capital stock and organizational knowledge, and whether to continue in business or to shutdown. If the firm exits, it receives a payment Φ_t for its remaining assets. Tomorrow's capital stock is certain, $K_{i,t+1} = K_{i,t}(1-\delta^K) + I_{i,t}^K$, whereas tomorrow's organizational knowledge is partly random and given by (3.2). Adjustment costs for organizational knowledge, $\psi^{\Omega}(I_{i,t}^{\Omega})^2/(2\Omega_{i,t})$, are quadratic as in a textbook model of Tobin's q. Similarly, adjustment costs for the capital stock are $\psi^K(I_{i,t}^K)^2/(2K_{i,t})$. Then the Bellman equation

[1]See equation 2.27 in chapter 2.

becomes

$$\begin{aligned} V(\Omega_{i,t}, K_{i,t}) \quad = \quad & \max \Bigg[\Phi_t, \sup_{I^\Omega_{i,t}, I^K_{i,t}, L_{i,t}} P^*(Z_{i,t}, \mathbf{D}_t)\, Z_{i,t} - w_t L_{i,t} - I^\Omega_{i,t} - I^K_{i,t} \\ & -\frac{\psi^\Omega}{2} \frac{(I^\Omega_{i,t})^2}{\Omega_{i,t}} - \frac{\psi^K}{2} \frac{(I^K_{i,t})^2}{K_{i,t}} + \frac{1}{R}\, \mathbb{E}\left[V(\Omega_{i,t+1}, K_{i,t+1})\, | \mathcal{F}_{i,t} \right] \Bigg], \end{aligned} \tag{3.6}$$

where $R \equiv 1+r$ is the real interest factor and $\mathcal{F}_{i,t}$ a firm's information set at time t. General market conditions, such as foreign market penetration, enter the decision through their effect on price. Each monopolist takes into account that higher supply depresses price given demand schedule (3.4). So, a monopolist sees price as a function $P^*(Z_{i,t}, \mathbf{D}_t)$, where $\mathbf{D}_t \equiv (N_t^{for}/N_t^{dom},\, \varepsilon_t, P_t^f, \tau_{i,t})$ stands for the vector of market conditions that firm i faces. Demand elasticity $-(1-\alpha)$ is constant, however, and independent of $\mathbf{D}_t$.

First, consider the case of a firm that continues in business. Tobin's q's for organizational knowledge and physical capital can be defined as

$$q^\Omega_{i,t} \equiv \mathbb{E}_t \left[\frac{1}{R} \frac{\partial V(\Omega_{i,t+1}, K_{i,t+1})}{\partial \Omega_{i,t+1}} \cdot x_{i,t+1} \right] \text{ and } q^K_{i,t} \equiv \mathbb{E}_t \left[\frac{1}{R} \frac{\partial V(\Omega_{i,t+1}, K_{i,t+1})}{\partial K_{i,t+1}} \right]. \tag{3.7}$$

Then, the first order conditions for the Bellman equation (3.6) imply that

$$q^\Omega_{i,t} = 1 + \psi^\Omega \frac{I^\Omega_{i,t}}{\Omega_{i,t}}, \quad q^K_{i,t} = 1 + \psi^K \frac{I^K_{i,t}}{K_{i,t}}, \quad \text{and} \quad L_t = L_0 + \frac{\alpha\beta}{w_t} P^*(Z_{i,t}, \mathbf{D}_t) Z_{i,t}. \tag{3.8}$$

Differentiating the value function with respect to the current state variable $\Omega_{i,t}$ and leading it by one period, one finds

$$R\, q^\Omega_{i,t} = \alpha\nu \mathbb{E}_t \left[\frac{P^*(\cdot)_{t+1} Z_{i,t+1}}{\Omega_{i,t+1}} \right] + \mathbb{E}_t \left[\frac{\psi^\Omega}{2} \frac{(I^\Omega_{i,t+1})^2}{(\Omega_{i,t+1})^2} \right] + (1-\delta^\Omega)\, \mathbb{E}_t \left[q^\Omega_{i,t+1} \right] \tag{3.9}$$

by (3.7) and the envelope theorem. An according condition applies to Tobin's q for physical capital.

So, under the usual regularity (no bubble) conditions,

$$q^\Omega_{i,t} = \frac{1}{1-\delta^\Omega} \sum_{s=t+1}^{\infty} \left(\frac{1-\delta^\Omega}{R} \right)^{s-t} \mathbb{E}_t \left[\frac{\nu}{\Omega_{i,s}}\, \alpha P^*(Z_{i,s}, \mathbf{D}_s)\, Z_{i,s} + \frac{\psi^\Omega}{2} \frac{(I^\Omega_{i,s})^2}{(\Omega_{i,s})^2} \right] \tag{3.10}$$

and

$$q^K_{i,t} = \frac{1}{1-\delta^K} \sum_{s=t+1}^{\infty} \left(\frac{1-\delta^K}{R} \right)^{s-t} \mathbb{E}_t \left[\frac{1-\beta}{K_{i,s}}\, \alpha P^*(Z_{i,s}, \mathbf{D}_s)\, Z_{i,s} + \frac{\psi^K}{2} \frac{(I^K_{i,s})^2}{(K_{i,s})^2} \right]. \tag{3.11}$$

A firm is uncertain about the realization of both future TFP and market conditions. The two terms in the expectations operator reflect the value of the respective state variable given market prospects $\alpha P^*(Z_{i,s}, \mathbf{D}_s)\, Z_{i,s}$ and savings in future adjustment costs $(I^K_{i,s})^2/(K_{i,s})^2$. So, market conditions affect the value of both state variables in a very similar way.

As a consequence, the model implies that a firm's capital stock and organizational knowledge are correlated from a researcher's perspective. By (3.8) and (3.2),

$$\Omega_{i,t+1} = x_{i,t+1} \cdot \Omega_{i,t} \left[\frac{q^{\Omega}_{i,t} - 1}{\psi^{\Omega}} + (1 - \delta^{\Omega}) \right]. \tag{3.12}$$

An according condition holds for $K_{i,t}$. So, for the researcher, the correlation between TFP and capital becomes

$$\mathbb{C}ov_t \left(\Omega_{i,t+1}, K_{i,t+1} \,|\Omega_{i,t}, K_{i,t} \right) = \frac{\Omega_{i,t} K_{i,t}}{\psi^{\Omega} \psi^{K}} \; \mathbb{C}ov_t \left(q^{\Omega}_{i,t} \, , \, q^{K}_{i,t} \right). \tag{3.13}$$

For the firm, $q^{\Omega}_{i,t}$ and $q^{K}_{i,t}$ are certain, given its information. The correlation is zero from its point of view. The researcher, on the other hand, does not know a firm's information set. Therefore, the data will exhibit a correlation between capital and TFP. The correlation is likely to be positive since future revenues affect both $q^{\Omega}_{i,t}$ and $q^{K}_{i,t}$ positively.[2] Olley and Pakes' original model does not allow for this possibility.[3]

However, since there is exit from the sample, the correlation in (3.13) does not give the

[2]Concretely, by (3.10) and (3.11),

$$\begin{aligned}(1-\delta^{\Omega}) q^{\Omega}_{i,t} \quad &= \quad (1-\delta^{K}) \rho_{i,t}\, q^{K}_{i,t} + \frac{1}{2} \sum_{s=t+1}^{\infty} \left(\psi^{\Omega} \left(\frac{1-\delta^{\Omega}}{R} \right)^{s-t} \mathbb{E}_t \left[\left(\frac{I^{\Omega}_{i,s}}{\Omega_{i,s}} \right)^2 \right] \right. \\ &\qquad \left. - \psi^{K} \rho_{i,t} \left(\frac{1-\delta^{K}}{R} \right)^{s-t} \mathbb{E}_t \left[\left(\frac{I^{K}_{i,s}}{K_{i,s}} \right)^2 \right] \right),\end{aligned}$$

where

$$\rho_{i,t} \equiv \frac{\nu \sum_{s=t+1}^{\infty} \left(\frac{1-\delta^{\Omega}}{R} \right)^{s-t} \mathbb{E}_t \left[\frac{P^*(\cdot)_s \, Z_{i,s}}{\Omega_{i,s}} \right]}{(1-\beta) \sum_{s=t+1}^{\infty} \left(\frac{1-\delta^{K}}{R} \right)^{s-t} \mathbb{E}_t \left[\frac{P^*(\cdot)_s \, Z_{i,s}}{K_{i,s}} \right]} > 0.$$

[3]It is sometimes argued that a positive productivity shock may push demand for a firm's good more than proportionally and thus capital input, giving rise to a positive correlation through demand rather than production effects. In a model of the present structure but with productivity beyond a firm's control, an exogenous productivity shock translates one to one into an output change with no effect on input choice. Specific assumptions on demand elasticity would allow a positive productivity shock to cause a more than proportional output increase and higher capital input. But *temporary* productivity shocks affect capital input little even under such assumptions.

complete picture. In general, the shutdown rule for a firm depends on the firm's state variables and its information about revenue prospects. Since the value function is increasing in both state variables, there are lower threshold levels for the states below which a firm exits, given market prospects. Alternatively, the shutdown rule can be written as a function of the realization of the TFP innovation. After observing the realization of $x_{i,t}$, a firm decides whether or not it prefers to exit. Then,

$$\chi_{i,t} = \begin{cases} 0 & \text{if } x_{i,t} < \underline{x}(\Omega_{i,t-1}, I^{\Omega}_{i,t-1}; K_{i,t}, \mathbf{D}_t) \\ 1 & \text{else} \end{cases} , \tag{3.14}$$

where $\chi_{i,t} = 0$ means that firm i chooses to shutdown at the beginning of period t. If the value of current and discounted future profits falls short of the outside value Φ_t, the firm has no incentive to produce in the current or any future period. Since the value function (3.6) is strictly increasing in the capital stock,[4] the threshold level $\underline{x}(\cdot)$ is strictly decreasing in $K_{i,t}$. A capital-rich firm is willing to bear lower TFP levels and still continues in business.

As Olley and Pakes (1996) pointed out, this introduces a negative correlation between the capital stock of survivors and the expected TFP level. Call the probability that a firm survives

$$Pr(\chi_{i,t+1} = 1 | \Omega_{i,t}, I^{\Omega}_{i,t}; K_{i,t+1}, \mathbf{D}_{t+1}) = P(\Omega_{i,t}, I^{\Omega}_{i,t}; K_{i,t+1}, \mathbf{D}_{t+1}). \tag{3.15}$$

Then by (3.2),

$$\mathbf{E}[\Omega_{i,t+1} | \chi_{i,t+1} = 1] = \left[(1-\delta^{\Omega})\Omega_{i,t} + I^{\Omega}_{i,t}\right] \int_{\underline{x}(\cdot)} x_{i,t+1} \frac{f\,(x_{i,t+1})}{P(\cdot)}\, dx_{i,t+1} \tag{3.16}$$

for the researcher. The firm is indifferent between staying in business and exit at the lower bound on $x_{i,t+1}$, $\underline{x}(\Omega_{i,t}, I^{\Omega}_{i,t}; K_{i,t+1}, \mathbf{D}_{t+1})$. The bound strictly decreases in the capital stock $K_{i,t+1}$. Thus, the value of the integral will be the lower the higher the capital stock happens to be. In the data, a negative relation between capital and the expected TFP level is likely to result. It is not clear *a priori* whether a positive correlation from (3.13) would outweigh the negative bias from (3.16) or

[4] $\partial V(\cdot)/\partial K_{i,t} = \alpha(1-\beta)P^*(\cdot)_t Z_{i,t}/K_{i,t} + (1-\delta^K)q^K_{i,t} + \psi^K (I^K_{i,t})^2/(2K_{i,t}) > 0$. Estimates in tables 3.6 and 3.9 confirm empirically that capital-rich firms are less likely to exit.

not. At any rate, a correlation exists between observed capital and unobserved TFP, and needs to be addressed in the estimation procedure.

3.3.3 Competition and a manager's efficiency choice

A maintained hypothesis of the present chapter is that trade liberalization induces managers to raise efficiency under fiercer international competition. Boone (2000), for example, shows conditions when more competition provides incentives to innovate products or processes. Hermalin (1992) and Schmidt (1997) present theoretical circumstances when increasing competition forces firms to reduce agency problems and managerial slack. An extension of the present model incorporates this insight and clarifies implications for productivity estimation.

A firm's investment in organizational skills $I^{\Omega}_{i,t}$ is observable to the firm's owner through cash flows. Similarly, $\Omega_{i,t}$ can be inferred from output. However, the manager's efforts in employing these organizational skills are not known. Successful efforts affect the distribution of the productivity shock $\tilde{x}_{i,t}$ in (3.2) favorably. In more profane words, efficiency improving investments are only successful if the management subsequently makes good use of the results. This gives rise to moral hazard. Suppose that a manager can either choose high efforts or low efforts ($E_{i,t} \in \{e^H_{i,t}, e^L_{i,t}\}$) and that the distribution of $x_{i,t+1}|e^H_{i,t}$ stochastically dominates the distribution $x_{i,t+1}|e^L_{i,t}$.[5] Under the assumption that efforts only affect next year's productivity $x_{i,t+1} \sim f(x_{i,t+1}|E_{i,t})$, it is easy to see that the firm's owner bases the optimal (end-of-year) remuneration $w(\cdot)$ on the observation of $x_{i,t+1}$. The owner maximizes $V(\Omega_{i,t}, K_{i,t}) - (1/R)\mathbb{E}\left[w(x_{i,t+1})\right]$ given the risk averse manager's participation constraint $\mathbb{E}\left[u\left(w(x_{i,t+1})\right)\right] - E_{i,t} \geq \underline{u}$ and the manager's optimality condition

$$\int_{\underline{x}} u\left(w(x_{i,t+1})\right) \frac{f(x_{i,t+1}|e^H_{i,t})}{1 - F(\underline{x}|e^H_{i,t})} \,\mathrm{d}x_{i,t+1} - e^H_{i,t} \geq$$

$$\geq \int_{\underline{x}} u\left(w(x_{i,t+1})\right) \frac{f(x_{i,t+1}|e^L_{i,t})}{1 - F(\underline{x}|e^L_{i,t})} \,\mathrm{d}x_{i,t+1} - e^L_{i,t}. \qquad (3.17)$$

[5] Second-order dominance is assumed for $\nu < 1$. A mean-preserving spread leaves the firm-fixed effect unchanged.

It is straight forward to use the principal's first order conditions and to show that the optimal pay for the manager is strictly increasing in $x_{i,t+1}$ if and only if the likelihood ratio $f(x_{i,t+1}|e^H_{i,t})/f(x_{i,t+1}|e^L_{i,t})$ is strictly increasing in $x_{i,t+1}$. Suppose this is the case.

Fiercer competition raises $\underline{x}(\cdot;\mathbf{D}_{t+1})$ and firms go out of business more frequently. Hermalin (1992) and Schmidt (1997) show that high-effort contracts can but need not become easier to institute under fiercer competition. A similar ambiguity arises here but for different reasons. Note that the likelihood ratio is increasing in $x_{i,t+1}$ and hence in $\underline{x}$, but the ratio $(1-F(\underline{x}|e^H_{i,t})/(1-F(\underline{x}|e^L_{i,t})$ is also increasing in $\underline{x}$. Together, these facts can but need not make the left-hand side in (3.17) grow larger relative to the right-hand side under fiercer competition and may facilitate the institution of high-effort contracts. Irrespective of whether competition has a positive or negative effect on efficiency, productivity estimation should account for it.

3.4 Firm-level Productivity

This section turns to the estimation of production functions at the level of sectors. Each firm's individual productivity is inferred. The estimation procedure is based on Olley and Pakes' (1996) algorithm to address the 'transmission bias' and extended to correct for the 'omitted price bias' in the spirit of Klette and Griliches (1996). The procedure also accounts for possible effects of competition on efficiency.

A 'transmission bias' in the capital coefficient arises because firms invest and decide whether to exit given their knowledge of productivity. Olley and Pakes' (1996) algorithm approximates the affected parts of the production function semiparametrically, uses the exit behavior of firms to extract additional information about expected productivity, and then removes the bias from the capital coefficients by leading the affected part of the production function one period. The remaining error term contains only an unpredictable shock to productivity, which is unrelated to a firm's investment choice.

When using revenues to approximate output, price is kept in the regression equation implicitly. An 'omitted price bias' arises. Productivity estimates are corrected for this following Klette and Griliches (1996). However, the structural model in the preceding section implies that the effect of expected revenues on TFP investment and the effect of competition on efficiency cannot be separated from the effect of omitted price. A remaining cyclical component in productivity estimates needs to be controlled for in subsequent analyses.

3.4.1 Production and foreign input efficiency

To measure the effect of foreign inputs on production directly, a modification of the production function is proposed here. Suppose firm i produces with the same Cobb-Douglas technology in every year t but with possibly different TFP. The capital stock is divided into three parts. Domestic and foreign equipment, $K^{dom}_{i,t}$ and $K^{for}_{i,t}$, and structures $S_{i,t}$. The variable structures includes real estate, premises, but also other capital goods such vehicles, computers, and rented or leased capital goods. Foreign equipment exceeds the efficiency of domestic equipment by a factor $(1+\gamma_K)$. The parameter γ_K is to be estimated and the hypothesis that $\gamma_K > 0$ will be tested.

Call the share of foreign equipment in total equipment $\kappa^f_{i,t} \equiv K^{for}_{i,t}/(K^{dom}_{i,t} + K^{for}_{i,t})$, the log of the total equipment stock with $k_{i,t}$, and the log of structures with $s_{i,t}$. Then, the production function can be written as

$$\begin{aligned} z_{i,t} &= \beta_K \ln\left(K^{dom}_{i,t} + (1+\gamma_K) K^{for}_{i,t}\right) + \beta_S\, s_{i,t} + \beta_L\, l_{i,t} + \beta_M\, m_{i,t} + \omega_{i,t} + \epsilon_{i,t} \\ &= \beta_K \ln\left(1 + \gamma_K\, \kappa^f_{i,t}\right) + \beta_K\, k_{i,t} + \beta_S\, s_{i,t} + \beta_M\, m_{i,t} + \beta_L l_{i,t} + \omega_{i,t} + \epsilon_{i,t} \\ &\approx \beta_K\, \gamma_K\, \kappa^f_{i,t} + \beta_K\, k_{i,t} + \beta_S\, s_{i,t} + \beta_M\, m_{i,t} + \beta_L\, l_{i,t} + \omega_{i,t} + \epsilon_{i,t}. \end{aligned} \tag{3.18}$$

$z_{i,t}$ denotes the logarithm of output, and $l_{i,t}$ denotes the log of the number of blue and white-collar workers (to be separated in the actual estimation). $m_{i,t}$ is the log of intermediate inputs. Since $\ln(1+c) \approx c$ for small values of c, one can recast the production function as on the third line, and a linear estimation technique can be employed.[6] The error term $\epsilon_{i,t}$ in (3.18) is a white noise shock

[6] Among the firms that dispose of foreign equipment, the average foreign equipment share is about 15.1 percent

Table 3.2: Select Production Function Estimates

Nív.50	Obs.	κ^f (1)	k (2)	s (3)	μ^f (4)	m (5)	l^{wh} (6)	l^{bl} (7)
08	*Machinery and equipment*							
	1,742	*-.037*	.047	.077	*.154*	.222	.238	.439
		(.129)	(.02)	(.022)	(.098)	(.014)	(.018)	(.025)
11	*Electronic equipment*							
	934	.24	.061	.053	*.094*	.244	.27	.261
		(.11)	(.022)	(.025)	(.121)	(.039)	(.024)	(.039)
12	*Cars, trucks, and buses*							
	308	*-.272*	.113	*.055*	*-.032*	.156	.186	.588
		(.179)	(.043)	(.038)	(.406)	(.035)	(.043)	(.061)
13	*Other vehicles and parts*							
	1,249	*.032*	.089	*.043*	.237	.221	.178	.532
		(.075)	(.019)	(.022)	(.094)	(.017)	(.019)	(.027)

κ^f: share of foreign equipment, k: log total equipment, s: log structures, μ^f: share of foreign intermediates, m: log total intermediates, l^{wh}: log number of white-collar workers, l^{bl}: log number of blue-collar workers.

Standard errors from 250 bootstraps. Estimates in italics not significant at the .95 level.

Data: *Pesquisa Industrial Anual* 1986-98, deflated with *IPA-OG*.

to the production technology, its variance (but not its mean) is taken to be constant across firms in a sector, and its realization is unknown both to a firm and the researcher.

The term $\beta_K \gamma_K \kappa^f_{i,t}$ measures the differential effect of foreign equipment on output. It can be interpreted as the efficiency difference between foreign and domestic equipment that would otherwise be attributed erroneously to TFP. A similar decomposition is made for the share of foreign intermediates μ^f in total intermediate inputs.

Production functions are estimated for 27 manufacturing sectors at *nível 50*, which corresponds to the SIC two-digit level. Table 3.2 gives an overview of the estimates for four select sectors that received much attention after trade reform. Table 3.10 at the end of this chapter lists all sectors and compares key estimates to fixed-effect regressions (FE), an alternative estimation method under the behavioral assumptions.

The equipment coefficients under FE are higher than the Olley-Pakes (OP) estimates in 21 out of 27 sectors, and the structures coefficients under FE are higher than the OP estimates in 10

in *PIA*. Among the firms that use foreign intermediates, the average share of foreign intermediates is 23.8 percent. Sample means are 3.1 and 10.3 percent, respectively. So, the approximation should be quite precise.

cases. So, a positive correlation between the productivity index and capital stocks occurs frequently. Pavcnik (2000) and Levinsohn and Petrin (2000) report larger ordinary least squares (OLS) than OP estimates for several sectors in Chilean industry. These findings cast doubt on the assumption that productivity evolves in a purely Markovian manner, whereas they can be explained by a q-theory model as in section 3.3.

To infer the efficiency differential of foreign equipment and intermediate goods, γ_K and γ_M, one can divide the coefficients in column 1 by the coefficients in column 2 and those in column 4 by column 5. The implied differentials are large in absolute value. For instance, $\gamma_K \approx 3.9$ for electronic equipment in table 3.2. The magnitude may be due in part to a bias from omitted variables such as managerial ability or quality and heterogeneity of output.

The discussion in section 3.5.1 (*Foreign Input Push*) will show that even high and potentially upward biased estimates for γ_K and γ_M do not yield a strong effect of foreign inputs on efficiency. The analysis is based on the more reliable measures $\beta_K \gamma_K$ and $\beta_M \gamma_M$. Foreign equipment is not always used more efficiently than domestic equipment. The coefficients on κ^f turn negative in 3 out of 7 sectors with significant estimates. Table 3.10 at the end of this chapter shows that γ_K varies between -8.6 and 15.7 (with a mean of 3.0) when significant, and γ_M takes values between .83 and 4.9 (mean 2.3) when significant. A negative coefficient can be interpreted as evidence that the average firm in a given sector fails to adjust its surrounding production process accordingly and does not realize the potential benefits of high-quality equipment. More on this in section 3.5.1.

3.4.2 Details on productivity estimation

This subsection discusses the precise estimation procedure in detail. Readers mostly interested in the effects of trade on productivity are encouraged to skip to section 3.5.

Production function (3.18) is augmented to account for all factors and estimated for 27 sectors under the restriction that all factor elasticities are constant between 1986 and 1998. This assumption yields time-invariant weights for the productivity indices. The variable κ^f is available for 1986 through 1995 and μ^f from 1996 to 1998. The observations are stacked accordingly so that $\beta_K \gamma_K$ and $\beta_M \gamma_M$ are identified in the respective subperiods.

To check for sensitivity, the data have been deflated with three different price indices. The sector-specific wholesale price index *IPA-OG* underlies all results in this chapter. Another sector-specific wholesale price index, *IPA-DI* (excluding imports), and the economy-wide price index *IGP-DI* (a combined wholesale and consumer price index) do not yield substantially different results. There is no producer price index for Brazil.

The productivity index $\omega_{i,t}$ follows from (3.2):

$$\omega_{i,t} = \nu \ln \left(\Omega_{i,t-1}(1-\delta^{\Omega}) + I^{\Omega}_{i,t-1} \right) + \beta_{0,i} + f(\mathbf{D}_t) + \xi_{i,t}, \tag{3.19}$$

where $\beta_{0,i} \equiv \nu \mathbb{E}\left[\ln x_{i,t}\right]$ is the firm-specific mean of productivity shocks, and $\xi_{i,t} \equiv \nu(\ln x_{i,t} - f(\mathbf{D}_t) - \mathbb{E}\left[\ln x_{i,t}\right])$ is a serially uncorrelated shock to productivity with mean zero and constant variance across firms in a sector. The function $f(\mathbf{D}_t)$ of market conditions captures their effect on the management's efficiency choice (see section 3.3.3). Both $\beta_{0,i}$ and $\xi_{i,t}$ are known to the firm when it chooses variable factor inputs and investment for next period. While entirely known to the firm's management, $\omega_{i,t}$ is unobservable to the researcher.

Correcting for 'Transmission Bias'

A transmission bias in the capital stock arises because both investment and the exit choice are correlated with $\omega_{i,t}$. A firm chooses organizational investment as a function of the state variables and market expectations. The model in section 3.3 implies that this choice is closely related to investment in capital goods. By (3.8), organizational investment is a function of q^{Ω}, $I^{\Omega}_{i,t-1} = (q^{\Omega}_{i,t-1} - 1)\Omega_{i,t-1}/\psi^{\Omega}$, and the q's for organizational skills and capital are positively related through (3.10) and (3.11): $q^{\Omega}_{i,t-1} = q(q^{K}_{i,t-1}; \cdot)$.[7] So,

$$I^{\Omega}_{i,t-1} = \frac{q(q^{K}_{i,t-1}; \cdot) - 1}{\psi^{\Omega}}\, \Omega_{i,t-1} = \frac{q(1 + \psi^{K} I^{K}_{i,t-1}/K_{i,t-1}; \cdot) - 1}{\psi^{\Omega}}\, \Omega_{i,t-1}.$$

[7] See footnote 2, p. 57.

As a consequence, the function $h(\cdot)$ in

$$\omega_{i,t} = h(I^K_{i,t-1}, I^S_{i,t-1}, a_{i,t}, k_{i,t}, s_{i,t}) + \beta_{0,i} + f(\mathbf{D}_t) + \xi_{i,t} \tag{3.20}$$

can be used to approximate the log TFP level in (3.19). In (3.20), $I^K_{i,t-1}$ and $I^S_{i,t-1}$ denote past investment in equipment and structures, respectively, and $a_{i,t}$ is the log of a firm's age at time t. Since market expectations affect the q's in a similar way, capital stocks and organizational skills should evolve roughly parallel to each other for surviving firms. A second-order polynomial in five variables related to business expectations (the foreign penetration rate, the real exchange rate, tariffs, aggregate sector-wide demand and the annual inflation rate) is used to proxy $f(\mathbf{D}_t)$. Interestingly, mostly labor coefficients change when $f(\mathbf{D}_t)$ is included. This points to a relationship between efficiency change and worker layoffs but is not further explored here.

Whereas Olley and Pakes (1996) have to assume that gross investment (capital good acquisitions less retirements) be strictly positive or at least non-zero to derive the equivalent of (3.20), no such assumption is needed in the present q-theory model. This is important. Firms stop to invest or even disinvest, especially in years before they go out of business. When excluding firms with zero gross investment, more than a third of all valid observations in the present sample would have to be ignored.[8]

The first regression equation follows,

$$\begin{aligned} z_{i,t} &= \beta_{0,i} + \beta_K\,\gamma_K\,\kappa^f_{i,t} + \beta_K\,k_{i,t} + \beta_S\,s_{i,t} + \beta_M\,\gamma_M\,\mu^f_{i,t} + \beta_M\,m_{i,t} \\ &\qquad + \beta_{bl}\,l^{bl}_{i,t} + \beta_{wh}\,l^{wh}_{i,t} + h(I^K_{i,t-1}, I^S_{i,t-1}, a_{i,t}, k_{i,t}, s_{i,t}) + f(\mathbf{D}_t) + \xi_{i,t} + \epsilon_{i,t} \\ &\equiv \beta_{0,i} + \beta_M\,\gamma_M\,\mu^f_{i,t} + \beta_M\,m_{i,t} + \beta_{bl}\,l^{bl}_{i,t} + \beta_{wh}\,l^{wh}_{i,t} \\ &\qquad + \phi(I^K_{i,t-1}, I^S_{i,t-1}, a_{i,t}, \kappa^f_{i,t}, k_{i,t}, s_{i,t}) + f(\mathbf{D}_t) + \xi_{i,t} + \epsilon_{i,t}. \end{aligned} \tag{3.21}$$

To be explicit, all regression variables are listed in (3.21). This includes the share of foreign intermediate inputs μ^f and the two groups of labor, blue- and white-collar workers. Only part of the

[8] More than 20,000 observations among the close to 60,000 valid ones exhibit zero gross investment in at least one capital good category, and 5,500 show negative gross investment.

equation is linear. The term $\phi(\cdot) \equiv \beta_K \, \gamma_K \, \kappa^f_{i,t} + \beta_K \, k_{i,t} + \beta_S \, s_{i,t} + h(\cdot)$ arises because the effect of log TFP on output cannot be separated from the effect of physical capital on output as long as their correlation is not removed. The coefficient estimates for β_{bl}, β_{wh} and β_M, on the other hand, are consistent if $\phi(\cdot)$ is approximated well. A polynomial series estimator of fourth degree is used here.[9] Neither year dummies nor time trend variables are significant when included. These findings lend support to the assertion that the drop in the sample in 1996 does not affect productivity estimates.

Next, one can estimate the probability of a firm's survival given today's information. This probability is given by (3.15) and can be restated as

$$Pr\,(\chi_{i,t} = 1|\cdot) = P(I^K_{i,t-1}, I^S_{i,t-1}, a_{i,t}, k_{i,t}, s_{i,t}; \mathbf{D}_t), \tag{3.22}$$

in the present context. There is no variable for a firm's expectations but one may suppose that the medium-sized to large firms in the sample are fairly well-informed about expected market outcomes. So, the vector of current market conditions $\mathbf{D}_t$, including tariff levels and market penetration rates, is taken as a proxy for past expectations. Equation (3.22) is the second estimation equation. It is estimated using a probit and a logit model.

Both the probit and the logit model predict slightly too few exits as compared to the data, and less dispersion. The inclusion of the vector of environment variables, $\mathbf{D}_t$, improves the correlation between probabilities (between zero and one) and observed outcomes (either zero or one) considerably. Financial variables of the firm turn out to reduce the fit and are not included. The logit model (correlation coefficient .223) outperforms probit (.211) in the estimation sample and is kept subsequently. Two different logit functions are estimated for the pre-1991 data and for the post-1991 data, taking into account that the shutdown probabilities may have changed systematically after trade liberalization. Contrary to the general finding that time indicators are not significant, the fit

[9] Levinsohn and Petrin (2000) argue that Olley and Pakes's (1996) algorithm suffers from two problems. First, they point out that investment in the past was made in anticipation of a firm-specific and forecasted shock to productivity. To account for this, equation (3.21) is estimated with the according fixed effect $\beta_{0,i}$ here. Second, Levinsohn and Petrin stress that a correlation between $\xi_{i,t}$ and the choice of labor and materials may exist, and address the issue in their algorithm. It follows from Newey (1994), however, that the estimates on current inputs are consistent under a proper series approximation. In fact, Levinsohn & Petrin find that only one in seven Olley & Pakes estimates for current inputs differs significantly (5% level) from the Levinsohn & Petrin estimates.

improves in this case.[10] A closer analysis of changes in exit (and other turnover) probabilities is provided in section 3.5.3.

Finally, to obtain a consistent estimate of the capital coefficients β_K and β_S, consider the contribution of capital to production one period in advance: $z_{i,t+1} - \beta_{0,i} - \beta_{bl}\, l^{bl}_{i,t+1} - \beta_{wh}\, l^{wh}_{i,t+1} - \beta_M\, \gamma_M\, \mu^f_{i,t+1} - \beta_M\, m_{i,t+1}$. Conditional on survival, the expectation of this term is

$$\begin{aligned}
&\mathbb{E}\left[z_{i,t+1} - \beta_{0,i} - \beta_{bl}\, l^{bl}_{i,t+1} - ... - \beta_M\, m_{i,t+1} \,\middle|\, A_{i,t}, \kappa^f_{i,t}, K_{i,t}, S_{i,t}; \omega_{i,t}, \chi_{i,t} = 1\right] \\
&\quad = \beta_K\, \gamma_K\, \kappa^f_{i,t+1} + \beta_K\, k_{i,t+1} + \beta_S\, s_{i,t+1} + \mathbb{E}\left[\omega_{i,t+1} \,|\omega_{i,t}, \chi_{i,t} = 1\right] \\
&\quad = \beta_K\, \gamma_K\, \kappa^f_{i,t+1} + \beta_K\, k_{i,t+1} + \beta_S\, s_{i,t+1} + \int_{\underline{\omega}(\cdot)} \omega_{i,t+1} \frac{f\left(\omega_{i,t+1}|h(\cdot)\right)}{P(\cdot)}\, d\omega_{i,t+1}
\end{aligned}$$

by equations (3.18), (3.20), and (3.22). Under regularity conditions (the density of $\xi_{i,t+1}$ needs to be positive in a neighborhood around $\xi_{i,t}$), $\underline{\omega}(\cdot)$ can be inverted and expressed as a function of $P(\cdot)$, too. So,

$$\begin{aligned}
&z_{i,t+1} - \beta_{0,i} - \beta_{bl}\, l^{bl}_{i,t+1} - ... - \beta_M\, m_{i,t+1} \\
&\quad = \beta_K\, \gamma_K\, \kappa^f_{i,t+1} + \beta_K\, k_{i,t+1} + \beta_S\, s_{i,t+1} \\
&\qquad + g\left(P(\cdot), \phi(\cdot) - \beta_K\, \gamma_K\, \kappa^f_{i,t} - \beta_K\, k_{i,t} - \beta_S\, s_{i,t}\right) + \xi_{i,t+1} + \epsilon_{i,t+1}
\end{aligned} \tag{3.23}$$

for some unspecified function $g(\cdot)$ since $h(\cdot) = \phi(\cdot) - (\beta_K\, \gamma_K\, \kappa^f_{i,t} + \beta_K\, k_{i,t} + \beta_S\, s_{i,t})$. $\xi_{i,t+1}$ is the unanticipated innovation in $\omega_{i,t+1}$. Hence, it is not correlated with net investment (gross investment less depreciation) or tomorrow's capital stock ($k_{i,t+1}$ and $s_{i,t+1}$), and the estimates of β_K, β_S and $\beta_K \gamma_K$ are consistent under the assumptions made. Equation (3.23) is the third estimation equation.

To approximate $g\left(P(\cdot), h(\cdot)\right)$ in equation (3.23), a third order polynomial expansion

$$\begin{aligned}
&z_{i,t+1} - \hat{\beta}_{0,i} - \hat{\beta}_{bl}\, l^{bl}_{i,t+1} - \hat{\beta}_{wh}\, l^{wh}_{i,t+1} - \hat{\beta}_M\, \hat{\gamma}_M\, \mu^f_{i,t+1} - \hat{\beta}_M\, m_{i,t+1} \\
&\quad = \beta_K \gamma_K \kappa^f_{i,t+1} + \beta_K k_{i,t+1} + \beta_S\, s_{i,t+1} + \sum_{m=0}^{3} \sum_{n=0}^{3-m} \beta_{m,n} (\hat{P})^m (\hat{h})^n + \eta_{i,t+1}
\end{aligned}$$

[10] No survival probability can be estimated for 1991 but is needed on the third step. In order not to lose all 1992 observations, the survival probability in 1991 is imputed as the unweighted average of the 1989, 1990, and 1992 predictions for each firm.

is used, where $\hat{\beta}_{0,i}$, $\hat{\beta}_{bl}$, $\hat{\beta}_{wh}$ and $\hat{\beta}_M$ are known from the first step. The capital coefficients enter this equation twice: in the additive terms, and through $\hat{h}(\cdot) = \hat{\phi}(\cdot) - (\beta_K \gamma_K \kappa^f_{i,t} + \beta_K k_{i,t} + \beta_S s_{i,t})$. The equation is estimated with non-linear least squares (NLLS), using the fixed-effects estimates of equation (3.18) as starting values. Subtracting the fixed effect $\beta_{0,i}$ from $z_{i,t}$ on the left hand side reduces the fit in some sectors. However, the error term needs to be identically distributed for the bootstrap to follow. This requires the subtraction of $\beta_{0,i}$.

Correcting for 'Omitted Price Bias'

While the production function is estimated consistently in section 3.4.2, a source of bias remains for productivity estimates. It arises because price is unknown but endogenous. Klette and Griliches (1996) address this problem.

The total of a firm's sales and production for store, deflated by sector-specific price indices, are used to approximate output. So, the dependent variable in the first regression equation (3.21) is in fact $p_{i,t} + z_{i,t} - \bar{p}_t$, where $p_{i,t}$ denotes the log of firm i's price and $\bar{p}_t$ the value of the price index used for deflation. By demand (3.3), the difference between a firm's price and market price is $p_{i,t} - \bar{p}_t = -(1-\alpha)d_{i,t} + (1-\alpha)(\bar{\theta}_t - \bar{p}_t)$, where $\bar{\theta}_t$ denotes the log of market-wide demand θY_t. Because of this relationship and since $d_{i,t} = z_{i,t}$ in equilibrium, the *de facto* regression is

$$\begin{aligned} p_{i,t} + z_{i,t} - \bar{p}_t &= \alpha z_{i,t} + (1-\alpha)(\bar{\theta}_t - \bar{p}_t) \\ &= \alpha\beta_{0,i} + (1-\alpha)\bar{\theta} \\ &\quad + \alpha\beta_M \gamma_M \mu^f_{i,t} + \alpha\beta_M m_{i,t} + \alpha\beta_{bl} l^{bl}_{i,t} + \alpha\beta_{wh} l^{wh}_{i,t} \\ &\quad + \alpha\phi(I^K_{i,t-1}, I^S_{i,t-1}, a_{i,t}, \kappa^f_{i,t}, k_{i,t}, s_{i,t}) + (1-\alpha)(\Delta\bar{\theta}_t - \bar{p}_t) \\ &\quad + \alpha\xi_{i,t} + \alpha\epsilon_{i,t}, \end{aligned} \tag{3.24}$$

instead of (3.21). Here, the log of market-wide demand for substitutes $(1-\alpha)(\bar{\theta}_t - \bar{p}_t)$ is decomposed into a preference based component $(1-\alpha)\bar{\theta}$ that does not vary over time, and into a time-varying component $(1-\alpha)(\Delta\bar{\theta}_t - \bar{p}_t)$ that moves with the business cycle ($\Delta\bar{\theta}_t \equiv \bar{\theta}_t - \bar{\theta}$).

The demand-side parameter α (which gives rise to a demand elasticity approximately equal to $-1/(1-\alpha)$) confounds the estimate of returns to scale by appearing in front of $z_{i,t}$. In addition, the time-invariant demand component $\bar{\theta}$ gets buried in the fixed-effects estimator. Klette and Griliches (1996) propose to use the sum of all firms' sales to approximate market-wide demand and to include it explicitly in the regression.Their purpose is to correct the scale estimate. Here, however, the focus lies on a consistent productivity estimate, and there are theoretical and practical reasons not to use Klette and Griliches' correction.

The model in section 3.3 (in particular (3.2) and (3.10)) implies that a firm's investment in $\omega_{i,t}$ depends on market expectations. In addition, the implementable efficiency choice of a manager depends on market conditions (3.17). If these market expectations are rational and firms are able to anticipate demand fairly well, the coefficient on sector-wide demand in an according regression will capture efficiency choice rather than the omitted price effect.

To understand the consequences, I also estimate (3.24) as suggested by Klette and Griliches (1996) and included the sum of sales (augmented by the degree of foreign market penetration) in the regression. The implied average log TFP level turns negative in all but two sectors. This finding indicates that market expectations go a long way in explaining productivity choice. Removing demand effects from productivity appears to be problematic. Moreover, production sectors may not coincide with consumer markets for the relevant substitutes. A wooden and a glass desk, for instance, show up in two different sectors in the present data but are close substitutes from a consumer's perspective. So, sector-wide sales seem to be too rough a proxy to demand θY_t in the relevant consumption markets.

Given these concerns, productivity estimates are only corrected for the time-invariant component $\bar{\theta}$. It can be extracted from the fixed-effects estimates by taking their sector-wide average. As a result, productivity estimates are clean of fixed demand components, but procyclical demand-side effects remain. To control for the remaining cyclicality, a demand proxy (the sum of sector-wide sales, augmented by the degree of foreign market penetration) will be included in all subsequent regressions. However, α will remain unidentified and no inference about economies of scale can be made. Yet, the assumption that α is constant across markets is likely not satisfied in practice so that economies of scale are not identified even if one is willing to make strong assumptions on the sources of cyclical TFP moves.

Estimates

Table 3.10 lists production function estimates for all 27 sectors and contrasts key estimates with fixed-effect regressions, an alternative estimation method under the behavioral assumptions. The fixed-effects correction under the OP method tends to reduce capital coefficients. In general, fixed-effects regressions of production functions are known to lower the capital coefficients. However, the behavioral model of section 3.3 favors this fixed-effects correction. On average across sectors, the sum of capital coefficients is about a quarter of the sum of labor coefficients. This is a low ratio. One might expect a ratio of double the magnitude. A reason for the low capital share may be that marginal returns on capital remain low despite low capital-labor ratios in Brazilian manufacturing or that production processes in Brazil are particularly labor intensive. Remaining measurement error in the capital stock series could bias the probability limits of the capital coefficients towards zero so that capital coefficients turn insignificant in some sectors. Since ratios rather than totals can be used to measure the effect of foreign inputs, formulation (3.18) makes sure that measurement error affects the estimates for $\beta_K \gamma_K$ and $\beta_M \gamma_M$ possibly little.

3.4.3 Total factor productivity

Given production function estimates, the logarithm of total factor productivity at the firm level is inferred as

$$
\begin{aligned}
\ln \mathrm{TFP}_{i,t} \quad = \quad & y_{i,t} - \widehat{(1-\alpha)}\bar{\theta} \\
& - \left(\hat{\beta}_K \, k_{i,t} + \hat{\beta}_S \, s_{i,t} + \hat{\beta}_M \, m_{i,t} + \hat{\beta}_{bl} \, l_{i,t}^{bl} + \hat{\beta}_{wh} \, l_{i,t}^{wh} \right),
\end{aligned}
$$

Data: Firm-level productivity in 27 manufacturing sectors in *Pesquisa Industrial Annual*.

Figure 3.3: Log TFP and labor productivity in manufacturing

where $y_{i,t} = (p_{i,t} - \bar{p}_t) + z_{i,t}$ denotes the total of deflated sales and production for store. The term $(1 - \alpha)\bar{\theta}$ corrects for fixed demand-side effects that affect productivity estimates through price $p_{i,t}$ in $y_{i,t}$ (Klette and Griliches 1996, see subsection 3.4.2). The quality of output and the number of varieties that multi-product firms produce are unobserved. As Melitz (2000) shows, both quality and variety increase the firm fixed effect $\beta_{0,i}$. Firm fixed effects are not subtracted from log TFP here, but will be controlled for subsequently. The efficiency contributions of foreign inputs $\hat{\beta}_K \hat{\gamma}_K \kappa^f$ and $\hat{\beta}_M \hat{\gamma}_M \mu^f$ are subtracted before the analysis of channels 2 and 3.

Figure 3.3 illustrates how TFP evolves in the aggregate of all 27 manufacturing sectors between 1986 and 1998. Except for a larger drop during the recession in the late eighties and the subsequent recovery, changes are small in general. At its trough, log TFP drops to .982 in 1990, but recovers and reaches 1.032 by 1998. Cavalcanti Ferreira and Rossi Júnior (1999) find a weaker recovery of TFP until 1997 to only about the level of 1986. Gomes (2001) reports similar, though more volatile aggregate TFP figures for Brazilian industry. The present study is the first to employ an extensive firm-level sample. Most previous studies on Brazilian industry considered labor productivity. As figure 3.3 shows, labor productivity increases more strongly than TFP during the 1990s because firms raise their capital stock.

3.5 Causes of Productivity Change

How does productivity change with trade liberalization? Do firms advance to best practice? If so, do foreign inputs contribute to the convergence? Do managers move their firms' efficiency ahead? Or does productivity improve primarily because the least competitive firms are shaken out? Questions like these can be related to three channels of trade effects on productivity: (1) A *Foreign Input Push* (section 3.5.1), (2) a *Competitive Push* (section 3.5.2), and (3) *Competitive Elimination* (section 3.5.3). An adequate way to evaluate the effects of trade on productivity seems to be a counterfactual approach. How would productivity have evolved in the absence of any of the three channels?

The present study treats foreign inputs as separate factors in the production function. Their effect on productivity is traced in subsection 3.5.1 (*Foreign Input Push*). Subsection 3.5.2 investigates whether reducing trade barriers has a positive effect on efficiency because of fiercer competition in the product market (*Competitive Push*). Subsection 3.5.3 analyzes to what degree inefficient firms are shaken out (*Competitive Elimination*) and sheds light on the question whether efficient firms become exporters. Subsection 3.5.4 discusses briefly the effects of potential further channels. Subsection 3.5.5 compares the three primary channels, posing the counterfactual that no trade liberalization was undertaken. The *Competitive Push* is singled out as the most important channel.

Table 3.3: Efficiency of Foreign Inputs

	11 *Electronics*		22 *Textiles*		24 *Leather*	
	log TFP	Input	log TFP	Input	log TFP	Input
	(1)	(2)	(3)	(4)	(5)	(6)
		$\beta_K\gamma_K\kappa^f$		$\beta_K\gamma_K\kappa^f$		$\beta_K\gamma_K\kappa^f$
1986	8.882	.004	7.564	.003	9.292	.004
1990	8.630	.024	7.522	.016	8.949	.018
1992	9.137	.027	7.611	.025	9.194	.031
1995	9.426	.085	7.498	.050	9.014	.065
		$\beta_M\gamma_M\mu^f$		$\beta_M\gamma_M\mu^f$		$\beta_M\gamma_M\mu^f$
1996	9.503	.044	7.460	.071	8.945	.204
1998	9.784	.047	7.500	.068	9.043	.199

Effect of foreign inputs in sectors with highest $\beta_K\gamma_K$ (24,11,22) estimates and highest $\beta_M\gamma_M$ (24) estimate.
Data: *Pesquisa Industrial Anual* 1986-98, deflated with *IPA-OG*.

3.5.1 Channel 1: Foreign Input Push

The counterfactual question for this channel is: How would firm productivity have evolved if firms had not been able to install foreign equipment, or if they had not been able to use foreign intermediates? Supposedly, foreign inputs exhibit higher quality and efficiency.

The estimated production functions have accounted for the potential efficiency differential. Under a fairly precise logarithmic approximation, the efficiency differential of foreign equipment and foreign intermediates *vis à vis* their domestic Brazilian counterparts is measured by $\beta_K\gamma_K\kappa^f$ and $\beta_M\gamma_M\mu^f$, where β_K and β_M are the elasticities of output with respect to total equipment and total intermediate goods. $(1+\gamma_K)$ and $(1+\gamma_M)$ are the factors of excess efficiency of foreign inputs, and κ^f and μ^f are the shares of foreign inputs in the respective totals. κ^f is available for the years 1986 through 1995, while μ^f is observed from 1996 until 1998. The coefficients are identified in an accordingly stacked system.

Table 3.3 summarizes mean log TFP and the effect of foreign inputs for the three sectors with the highest significant $\beta_K\gamma_K$ estimates (electronic equipment, textiles, and leather products and footwear). The leather and footwear sector exhibits the highest $\beta_M\gamma_M$ estimate. The figures show that foreign input efficiency contributes only minimally to productivity. It is orders of magnitude smaller than average productivity even in the sectors with the strongest differentials. Foreign input efficiency neither serves as a break in times of falling productivity nor as a push in times of rising log TFP. Take the electronics sector as an example. Between 1986 and 1990, firms invested strongly in foreign equipment and pushed $\beta_K \gamma_K \kappa^f$ from .004 to .024. Without that push, log TFP would have fallen to 8.61 but foreign equipment stopped the fall at 8.63. This is less than a one percent difference for one of the strongest positive effects in the sample.

Some sectors exhibit significantly negative estimates of γ_K. They can be viewed as evidence that the average firm in the sector fails to effectively implement foreign equipment in the short term. Technology adaption takes time because of factor complementarities, learning effects and necessary production rearrangements. Similar arguments have been advanced to explain the productivity slowdown in industrialized countries in periods of technology adoption. Foreign machines of high quality sell at a price premium over domestic counterparts, and firms need to put foreign machines to more efficient uses than domestic ones in order to avoid a productivity loss. To test for this supposition, one can split κ^f into recent-year investment and the lagged κ^f level and re-estimate production. Two of the three sectors with significantly negative κ^f estimates confirm that the recent-year term exceeds the longer-back foreign equipment term in absolute value and can explain most of the negative effect. Delayed adjustments to new machinery may thus be behind both occasional negative effects of foreign machinery on productivity and the small positive effects.

It has been argued that firms may benefit from embodied technology when acquiring foreign goods. That is, foreign drilling machines or turning lathes are supposed to do more than just process a workpiece. They are thought to be essentially different from their domestic counterparts under this hypothesis. If it is true, foreign inputs should enter the production function separately and interact with other factors in a different way than domestic inputs. However, foreign inputs are often zero. In fact, 79.7 percent of all firms in 1986-1995 dispose of no foreign machines, and 56.8 percent of all firms in 1996-1998 use no foreign intermediate inputs. So, standard production functions cannot be estimated. Instead, a Box-Cox transformation can be used for both types of foreign inputs. Production functions were re-estimated under an accordingly adjusted OP procedure.

With the Box-Cox transformation, TFP can be reassessed under the extreme counterfactual hypothesis that all inputs had to be Brazilian rather than partly foreign. In the case of foreign equipment, for instance, the difference $[\hat{\beta}_{K^f}((K^f_{i,t})^{\hat{\lambda}_K} - 1)/\hat{\lambda}_K + \hat{\beta}_{K^d} \ln K^d_{i,t}] - [\hat{\beta}_{K^d} \ln(K^d_{i,t} + K^f_{i,t})]$ can be taken as a measure for the contribution of foreign equipment efficiency. It can be compared to the value in columns 2, 4 and 6 of table 3.3 and should be understood as the difference that setting κ^f to zero would make. Resulting log TFP figures are lower and behave more erratically under a Box-Cox transformation, while estimates of input efficiency differentials are higher. However, the relative magnitude of foreign input efficiency is still not high enough to account for much TFP change over time. This corroborates previous findings and there is little evidence that effects of embodied technology are sources of immediate productivity change.

Keller (2000) finds for a sample of industries in 8 OECD countries that machinery imports matter but that their impact may be limited conditional on the effect of domestic technology. To my knowledge, only Feenstra, Markusen and Zeile (1992) and Fernandes (2001) estimate the effect of inputs on production at the micro-level. Feenstra et al. (1992) distinguish the effect of more inputs of the same type from the effect of a greater range of them in a sample of Korean *chaebol*—albeit not with respect to foreign trade. They detect a positive correlation between their input measure and the change in TFP. Using a large sample of Colombian plants, Fernandes (2001) finds that productivity gains are stronger in sectors that use foreign intermediates to a higher degree. However, neither one of the studies reports how much TFP change their estimates predict and their findings cannot be compared to those of table 3.3.

Related studies for developed countries corroborate these findings. Black and Lynch (2001), for instance, analyze a representative sample of businesses in the USA. They find that productivity improvements do not depend on whether an employer adopts new work practices but rather how an innovative work practice is implemented within the establishment. Funk and Strauss (2000) show that productivity change triggers (Granger-causes) capital investment, and not the other way around. Their analysis employs a time-series of investment in 450 US manufacturing sectors. Results of the present chapter paint a similar picture. The causal link from the quality of installed equipment to productivity does not appear to be strong. To make appropriate use of new inputs, firms need to embed foreign equipment into the production process and may have to adopt new processes. If they can take such measures only over time, foreign inputs will not create value beyond cost in the short run.

3.5.2 Channel 2: Competitive Push

Increased foreign competition can foster product and process innovation (Boone 2000). Foreign competition may also end the *quiet life* of managers and allow firms to enforce higher efficiency (Hermalin 1992, Schmidt 1997). The counterfactual question is: What would firm-level productivity have looked like had there not been an increase in competitive pressure due to foreign imports, or the threat of more foreign imports? To find an answer, the change in firm-level productivity can be regressed on two variables related to foreign competition: the nominal tariff in the firms' respective output markets and the penetration of their markets with foreign imports. Market penetration proxies the level of non-tariff barriers in Brazil, while nominal tariff levels capture the effect of tariff barriers directly. Foreign penetration is measured as the share of imports per absorption in a given market.

However, there are econometric concerns. Market penetration and low tariff barriers may not only induce firms to strive for higher productivity. The causation can also go in the opposite direction. Consider tariffs. When motivating trade reform, the Brazilian government declared that it aimed to instill efficiency change through foreign competitive pressure. If the government obeyed the principle, it applied lower tariffs to low-growth sectors. This introduces a positive correlation between TFP change and tariff levels. Second, take market penetration. When barriers to imports fall, the least efficient sectors are likely to attract the strongest influx of competing imports. In other words, low productivity performance may cause high market penetration, which brings about a negative correlation between TFP change and market penetration.

Instrumental variables can remedy both sources of simultaneity. Foreign market penetration not only depends on tariffs and competitors' productivity but also responds to a country's terms of trade. The real exchange rate fluctuated considerably over the period 1986 to 1998 and was thus an important factor for the relative price of imports. Certain components of the real exchange rate are exogenous variables in the sense that they affect foreign firms' entry decision (and the government's tariff choice) but are unanticipated by Brazilian firms.

The real exchange rate is decomposed here into the nominal exchange rate relative to the US dollar and into two foreign price components, an average European and an average Asian price index (using Brazilian imports in 1995 as weights). Nominal exchange rates are difficult to predict, and Brazilian firms were most likely not able to anticipate the US dollar exchange rate. This makes the nominal exchange rate a valid instrument. Brazilian domestic inflation was more predictable and had likely an impact on managers' efficiency choice. It is therefore not considered a valid instrumental variable, notwithstanding its importance for the real exchange rate. One might suspect that Brazilian manufacturers were able to anticipate well the price level of major trading partners such as the US (number one in 1995) or Argentina (number two). Firms can expect less foreign competition whenever a high price level of a major trading partner raises the bilateral real exchange rate, and vice versa. The suspicion is testable. It turns out that the US price index is a valid instrument, but not Argentina's price level.

On the other hand, the average European price index and the average Asian price index are taken *a priori* to be valid instruments for two reasons. First, the exports of foreign firms from these regions appear to be less sensitive to price movements. Most countries in these regions are small trading partners of Brazil, and the quantity of specialized products they sell to Brazil is likely to remain about the same even if prices fluctuate. Second, Brazilian firms are less likely to anticipate correctly the effect of prices in regions that are not major trading partners. The later result that even the US price index is a valid instrument justifies this reasoning *ex post*. There are two endogenous variables, tariffs and market penetration, and three instruments, the nominal exchange rate, the European and the Asian price index. When dropping the European price index, a Hausman test for overidentification shows no evidence that the European price index is not a valid instrument.

Table 3.4 summarizes the regression results under both a two-stage least squares approach (columns 1 and 5) and a simple regression (column 4). The dependent variable is the first difference in log TFP (except for the first-stage IV regressions in columns 2 and 3). Unobserved firm-specific factors such as managerial ability are likely to affect a firm's TFP level. So, a fixed-effects model is employed throughout and standard errors are corrected accordingly.

Both tariffs and market penetration are measured on a scale from zero to one. Lower tariffs induce firms to raise efficiency, as does higher market penetration (column 1). The effects are considerable. A reduction of nominal tariffs by 10 percentage points (.1) induces firms to increase log TFP by .225. An increase of foreign market penetration by 1 percentage point (.01) raises TFP by another .28. Log TFP is about 8.1 on average across all sectors and years. So, a reduction of nominal tariffs by 10 percentage points pushes log TFP by 2.8 percent ($[.225/8.1] * 100$). A careful counterfactual simulation will follow in section 3.5.5.

How important were the treatments in productivity estimation for this result? Column 5 shows that estimates would be slightly less favorable but not strongly different when inferring productivity from simple OLS regressions on the unbalanced panel. Piecewise estimations reveal that the point estimates change mostly because of the correction following from section 3.3.3 and little because of Olley and Pakes. To isolate the effect of foreign competition from possibly confounding effects, firm-level variables such as foreign inputs and indicators for firm size are included in the regression. Firms that start to use more foreign inputs suffer a slowdown in productivity in the subsequent year. They face implementation costs and may need to train workers and carry out adjustments to the production process (subsection 3.5.1).

The stock of sector-wide invested foreign capital correlates positively with productivity

Table 3.4: Foreign Competition and Productivity Change

	2SLS-FE			FE	2SLS-FE
	Δ ln TFP	Tariff	M.Pen.	Δ ln TFP	OLS Δ ln TFP
	(1)	(2)	(3)	(4)	(5)
Nominal tariff	-2.249			-.115	-1.467
	(.201)			(.027)	(.186)
Market penetration	27.989			.691	23.056
	(2.163)			(.151)	(2.006)
Age	-.294	-.022	.008	-.02	-.22
	(.023)	(.0007)	(.0001)	(.002)	(.021)
κ^f	-.424	-.058	.008	-.092	-.191
	(.053)	(.006)	(.001)	(.029)	(.049)
μ^f	-.504	.05	.027	.173	-.477
	(.074)	(.007)	(.001)	(.033)	(.069)
ι(medium L^{tot})[a]	.116	*.0003*	*.0007*	.123	.113
	(.043)	(.006)	(.001)	(.029)	(.04)
ι(big L^{tot})[a]	.176	*-.007*	*-.0006*	.152	.175
	(.045)	(.006)	(.001)	(.03)	(.042)
ι(medium cap.)[b]	-.038	.018	.001	-.038	*-.03*
	(.017)	(.002)	(.0004)	(.011)	(.016)
ι(big cap.)[b]	*-.017*	.024	*.0002*	-.047	*-.007*
	(.024)	(.003)	(.0006)	(.015)	(.022)
Sector demand[c]	-.548	-.02	.012	-.225	-.455
	(.034)	(.003)	(.0005)	(.013)	(.031)
Cum. FDI[d]	.185	.027	-.004	-.014	.14
	(.019)	(.001)	(.0003)	(.007)	(.018)
FDI flow[d]	-.14	-.034	.0008	-.029	-.098
	(.016)	(.002)	(.0003)	(.008)	(.015)
Nom. exch. rate (USD)		.04	.035		
		(.01)	(.002)		
Prices EU		-.587	-.015		
		(.013)	(.002)		
Prices Asia		*.041*	.171		
		(.062)	(.012)		
Obs.	33,493	33,493	33,493	33,493	33,493
R^2	.0002	.846	.551	.0005	.0001

[a] Medium: ($30 \leq L_{i,t}^{tot} < 300$), big: ($L_{i,t}^{tot} \geq 300$).

[b] Medium: $K_{i,t} + S_{i,t}$ in middle tercile of all firms in a year, big: in upper tercile.

[c] Sector-wide sales in *PIA*, augmented by foreign market penetration.

[d] Billion USD per sector. Cumulated FDI is end-of-year stock of invested foreign capital.

increases at the firm-level. Foreign direct investment (FDI) directed to a sector as a whole forces each individual firm to improve efficiency because foreign-owned domestic competitors are likely to become more productive with foreign capital. So, FDI may work like a substitute for trade liberalization. However, it takes USD 1 Billion to raise log TFP by .185—an increase that a tariff reduction by 8 percentage points (.08) can also achieve. The invested foreign capital stock in Brazilian manufacturing totalled USD 30 Billion in 1998. In this light, an FDI inflow of USD 1 Billion in a single sector would be substantial. In addition, FDI flows seem to have an offsetting negative effect on productivity. This may be because FDI to foreign-owned domestic competitors also reduces the market penetration of foreign firms that export to Brazil (column 3).

In the first stage of the instrumental variable estimation (columns 2 and 3), fixed-effects regressions are run using all observations. This makes the regressions in columns 2 and 3 weighted ones. The tariff cannot be used as a predictor of market penetration (column 3). If included, order conditions would fail and the system would not be identified. Separate regressions show that market penetration falls 1.73 percentage points on average across all sectors when tariffs are raised 10 percentage points.

The consistency of estimates in table 3.4 depends on the validity of the proposed instruments. A comparison of a simple regression (column 4) to the current results (column 1) confirms that the suspected negative bias in tariffs and the positive bias in the market penetration do exist. The same biases can be detected in a level regression, as table 3.8 at the end of this chapter shows (column 8 vs. 5). In this case, the bias is strong enough to reverse the sign on tariffs. Table 3.8 at the end of this chapter also documents the procedure to test for the validity of instruments. Depart from a regression that includes only the clean instruments (the nominal US dollar exchange rate, the European Union's (EU) price level and the Asian price level; columns 1 and 5). Additional instruments can be inserted and tested using a Hausman test for overidentification. For the US price level, the test statistic is negative in the TFP change regression (column 3) and the test cannot be conducted. However, the US price level turns out to be a valid instrument in the TFP level

Table 3.5: Transitions Between States of Operation Before and After 1991

Firm $\sigma_{i,t}$ \ $\sigma_{i,t+1}$	*active* *exporter* (1)	*active* *non-exporter* (2)	*suspended* (3)	*extinct* (4)	*Total*
active					
exporter	86.2 ▷ 88.2	12.5 ▷ 8.5	.8 ▷ .9	.6 ▷ 2.4	*100.0*
non-exporter	3.7 ▷ 7.0	91.2 ▷ 85.9	1.6 ▷ 2.0	2.8 ▷ 5.1	*100.0*
suspended	1.9 ▷ 7.6	31.6 ▷ 32.5	57.3 ▷ 40.4	9.2 ▷ 19.5	*100.0*
extinct	.0 ▷ .0	.0 ▷ .0	.0 ▷ .0	100.0 ▷ 100.0	*100.0*

$Pr(\sigma_{i,t+1}|\sigma_{i,t})$

Source: Own calculations from observed transitions. (Observations of mergers, acquisitions and split-ups treated as missing. Transitions 1990-92 treated as if 1990-91.)

Data: *Pesquisa Industrial Anual*. Firm-level exports from SECEX, 1989-1998.

regression (column 6). This can be interpreted as *ex post* evidence that the European and Asian price levels are also valid. When dropping the European price index, an overidentification test shows no evidence that the European price index is invalid (not reported). The price indices of Brazil and Argentina, however, are not valid instruments (columns 2, 4 and 7).

Only regressions of the *change* in log TFP separate the present channel 2 from effects of exit from the sample (channel 3). A difference-in-difference analysis confirms that exiting firms have lower productivity on average. Fiercer foreign competition is likely to bring about more exits. So, level regression would confuse the two channels and inappropriately boost the estimates.

3.5.3 Channel 3: Competitive Elimination

What would industry turnover have looked like in the absence of trade liberalization? There are many aspects to industry turnover and it is hard to link them directly to the trade regime. A new method to evaluate turnover causally is proposed here: the estimation of Markov probabilities for an active firm's transition between possible states of operation.

The transition probabilities in table 3.5 reflect the likely pattern of a Brazilian manufacturer's choice between 1989 and 1991 (to the left of the arrow), and between 1991 and 1998 (to the right of the arrow). Data on the exporting status of firms are not available before 1989. The competitive environment changes with trade liberalization and the transition matrices reflect this. The survival probability of a non-exporter, for instance, drops from 95.6 percent to 92.9 percent (100 percent less estimates in columns 3 and 4).

To evaluate causally how the trade regime influences turnover, transition probabilities can be estimated as functions of the market environment and firm characteristics, among them productivity. Unnested unconditional multinomial logit (MNL) appears to be an appropriate estimation technique because the states of operation have no specific characteristics and independence from irrelevant alternatives holds. A firm's probabilistic choice is taken to satisfy

$$Pr(\sigma_{i,t+1}|\sigma_{i,t};\mathbf{x},\beta_{\mathbb{A}}) = \frac{e^{\beta'_{\sigma}\mathbf{x}}}{\sum_{\varsigma\in\mathbb{A}} e^{\beta'_{\varsigma}\mathbf{x}}},$$

where the choice set $\mathbb{A}$ includes four alternatives for $\sigma_{i,t+1}$: to be an exporter, to be a domestically active firm only, to suspend production temporarily, or to exit. The model is estimated independently for the three possible current states $\sigma_{i,t}$: exporter, non-exporter, or temporarily suspended firm.

Table 3.6 reports results for active firms ($\sigma_{i,t}$: exporter or non-exporter), and table 3.9 at the end of this chapter presents the remaining category ($\sigma_{i,t}$: suspended firm). Since probabilities have to sum to unity, the parameter vector β_σ is only identified for three choices relative to a fourth choice of reference. Here, the current states of operation ($\sigma_{i,t+1} = \sigma_{i,t}$) are chosen as the respective points of reference. The reference for a non-exporter, for instance, is that the firm remains a non-exporter.

In order to look at the effect of trade on turnover beyond previous channels, a firm's TFP needs to be controlled for. Productivity itself has the expected effect on turnover. The lower it is, the more likely a firm exits or suspends production (columns 2, 3, 5 and 6 in table 3.6). Both theoretical and empirical evidence suggests that more efficient firms self-select into becoming exporters (Hopenhayn 1992, Melitz 1999, Clerides, Lach and Tybout 1998, Bernard and Jensen 1999). The present analysis supports this hypothesis. When productivity is high, non-exporters start exporting more often (column 4) and exporters abandon exporting less frequently (column 1).

Table 3.6: Multinomial Logit Estimates of Transition Probabilities

$\sigma_{i,t}$	Exporter			Non-Exporter		
$\sigma_{i,t+1}$	Non-Exp.	Susp.	Exit	Exp.	Susp.	Exit
	(1)	(2)	(3)	(4)	(5)	(6)
Tariff[a]	2.026	*-.507*	-1.915	*-.187*	-1.028	-1.049
	(.288)	(.962)	(.855)	(.139)	(.46)	(.373)
Real exch. rate (USD)[b]	-.728	*.262*	-2.84	-1.454	*-.455*	-3.369
	(.247)	(.722)	(.59)	(.115)	(.375)	(.322)
ln TFP	-.078	-.214	*-.003*	.072	-.28	-.232
	(.035)	(.083)	(.066)	(.019)	(.057)	(.04)
Age	*0*	*-.004*	*.003*	*.0002*	*-.003*	*.005*
	(.002)	(.007)	(.004)	(.001)	(.005)	(.003)
κ^f	-.752	*-1.014*	-2.187	1.239	*-.122*	*-.658*
	(.296)	(1.137)	(.822)	(.213)	(.84)	(.58)
μ^f	-1.022	*-.117*	*.096*	*.102*	*-3.494*	*-.739*
	(.389)	(.773)	(.517)	(.344)	(2.628)	(.762)
ι(med. L^{tot})[c]	-1.01	*-1.045*	-1.649	.457	*-.29*	-.803
	(.345)	(.746)	(.523)	(.171)	(.264)	(.16)
ι(big L^{tot})[c]	-1.612	-2.169	-1.887	1.094	-.891	-1.196
	(.347)	(.761)	(.532)	(.173)	(.303)	(.192)
ι(med. cap.)[d]	-.663	*.459*	-1.1	.558	*-.034*	*-.148*
	(.19)	(.708)	(.335)	(.109)	(.193)	(.129)
ι(big cap.)[d]	-1.079	*.565*	-1.436	1.239	*.009*	*-.147*
	(.201)	(.726)	(.357)	(.115)	(.249)	(.171)
Sector demand[e]	*.041*	*-.068*	.121	-.112	*-.025*	.14
	(.023)	(.057)	(.039)	(.01)	(.024)	(.018)
Cum. FDI[f]	.097	-.215	-.215	*.027*	*-.032*	*-.032*
	(.037)	(.067)	(.067)	(.025)	(.043)	(.043)
FDI flow[f]	*.03*	.642	.342	-.426	*.162*	*-.123*
	(.092)	(.227)	(.168)	(.065)	(.176)	(.116)
Obs.		13,123			27,424	
Pseudo R^2		0.694			0.664	
$\hat{\chi}^2$		8412.43			22780.80	
$Pr(\chi^2_{39} > \hat{\chi}^2)$		.0000			.0000	

[a] Next year's tariff.

[b] Annual. Based on *IPA-OG* and US producer price index.

[c] Medium: $(30 \leq L^{tot}_{i,t} < 300)$, big: $(L^{tot}_{i,t} \geq 300)$.

[d] Medium: $K_{i,t} + S_{i,t}$ in middle tercile of all firms in a year, big: in upper tercile.

[e] Sector-wide sales in *PIA*, augmented by foreign market penetration.

[f] Billion USD per sector. Cumulated FDI is end-of-year stock of invested foreign capital.

Incentives for exporting from Brazil hardly changed over the period. So, one may interpret the positive effect of higher productivity as close to causal: Higher productivity encourages firms to be exporters.

Findings for both tariffs and the real exchange rate show that reduced barriers to imports cause more exits. Firms choose next period's state of operation with regard to market prospects (exit in the data means exit in the following year). So, tariffs here are next year's tariffs. The lower the tariff, the more likely it is that a firm goes out of business (columns 3 and 6). The estimate of -1.049 in column 6 means that a reduction of tariffs by 10 percentage points (.1) raises the exit probability by 1.11 ($= e^{.1049}$) percent relative to a non-exporter's likelihood of remaining a non-exporter. Similarly, lower tariffs make it more likely that a firm suspends production (column 5), possibly to wait for a return to higher tariff protection.

A low real exchange rate has a similarly strong effect on exit (columns 3 and 6) but no significant effect on the suspension decision (columns 2 and 5). Since firms are likely not able to predict the real exchange rate, current levels are used in the regression. The lower the real exchange rate, the harder it is to compete abroad, and more Brazilian exporters stop exporting (column 1). Interestingly, a low real exchange rate induces non-exporting firms to start exporting (column 4). The result could imply that exporters benefit from observing the influx of foreign goods to identify internationally competitive product characteristics.

In the previous MNL regressions, both the government's choice of tariff levels and the real exchange rate are taken as exogenous to transition choices. The Brazilian government aimed to induce a competitive push. The inclusion of log TFP in the regression controls for this. Only if the Brazilian government also wanted to induce more exits with trade reform, a simultaneity problem would arise. If existent, the simultaneity is alleviated because future rather than current tariffs are used in the regression. To check the estimates nevertheless, the same instrumental variables as in section 3.5.2 can be used to predict tariffs. The non-linear structure of the logit regression does not allow to remove the simultaneity if existent. However, a regression with predicted variables can serve as a robustness check. If coefficient estimates do not change significantly, simultaneity is likely not a problem. Table 3.9 at the end of this chapter reports estimates when predicted tariffs are used. The estimates hardly change compared to the previous regression. This is especially the case for the exit decision, which is most important here.

A difference-in-difference analysis shows that exiting firms have 8.2 percent lower productivity than survivors on average. So, exits may help raise average productivity. However, the shutdown probability ranges between two and five percent only. The bearing of exits on aggregate productivity remains to be evaluated. A counterfactual simulation follows in section 3.5.5.

3.5.4 Possible additional channels

Entry is another aspect of turnover. Fiercer foreign competition can deter entry—a competitive elimination of business projects before they are realized. However, entry is excluded from the present analysis for two reasons. For one, entry was not always recorded systematically in *PIA*. Second, the counterfactual is hard to answer in general: How many more business proposals would have been pulled out from the drawers had trade not been reformed? It is likely that only the most productive projects will be realized after trade reform. Then the net effect on efficiency is ambiguous. Less but more productive entrants can move aggregate productivity either way.

At least from a theoretical perspective, there are two additional channels through which trade may affect productivity. In the aggregate of sectors, a fourth channel can be *Induced Size Change*. Less competitive firms lose market share, while more competitive firms grow in relative size. Models with Cournot or monopolistic competition predict this. The effect raises sector-wide productivity because averages are size-weighted. It is difficult, however, to causally relate size change to trade liberalization. In fact, it is likely to be an indirect effect in several ways. First, trade encourages firms to raise individual productivity through the foreign input push and the competitive push. Firms that are faster at adopting higher productivity grow in relative size. Therefore, size change gives the foreign input push (1) and the competitive push (2) an extra boost. Similarly, after suspension or exit has occurred due to competitive elimination (3), the surviving firms grow in size, and the fittest among them grow relatively faster. In this way, size change also gives channel 3 an extra boost. Finally, increased foreign competition squeezes the market share of domestic Brazilian firms. Again, the less productive ones are squeezed more strongly which promotes channels 1 and 2. This suggests that, on all of these accounts, size change should not necessarily be considered its own channel but rather an augment to previous channels.

Size change does seem to be a channel of its own with regard to economies of scale. If economies of scale exist, firms that face import competition may suffer from lower scales of production after being squeezed, while exit of their domestic competitors helps them realize previously unexploited economies of scale. So, there are conflicting forces at work and it is not clear which would prevail. Studies that investigate scale effects from trade are, in general, not able to confirm an effect empirically (Tybout and Westbrook 1995, Roberts and Tybout, eds 1996). Unfortunately, productivity and economies of scale are not identified simultaneously (see section 3.4.2). So, this channel cannot be evaluated in the present context.

Regarding trade effects in the aggregate of industry, a fifth channel of trade is *Induced Specialization*. Due to Ricardian or HOV type forces of trade, a country's industry may specialize in sectors where the innovative potential is largely exhausted. This can lower average productivity of industry as a whole. Theoretical contributions in favor of this hypothesis include Young (1991) and Xie (1999). I have expressed theoretical doubt in chapter 2. This fifth channel seems to be difficult to evaluate in general and especially so in the current context with incomplete sector data. Using cross-country data, Weinhold and Rauch (1999) find empirical evidence against the hypothesis.

3.5.5 Counterfactual simulations

The first row in table 3.7 shows how productivity evolves in the sample. To assess the relative importance of the three channels, one can switch them off individually and calculate log TFP in their absence. Trade reform took effect in 1990, whereas previous tariff reductions most

Table 3.7: Counterfactual Simulations

	Counterfactual	log TFP 1986	1990	1992	1995	1998
De facto		1.0	.9821	.9983	1.0051	1.0317
Ch. 1 off	κ^f and μ^f lower[a]	1.0	.9776	.9979	1.0050	1.0312
Ch. 2 off	Tariffs unchanged[b]	1.0	.9679	.9604	.9976	1.0356
Ch. 3 off	Tariffs unchanged	1.0	.9826	.9979	1.0034	1.0278

[a]Based on separate regressions, a one percentage point lower tariff is assumed to result in a 28.9 percentage point higher demand for foreign inputs relative to domestic inputs.

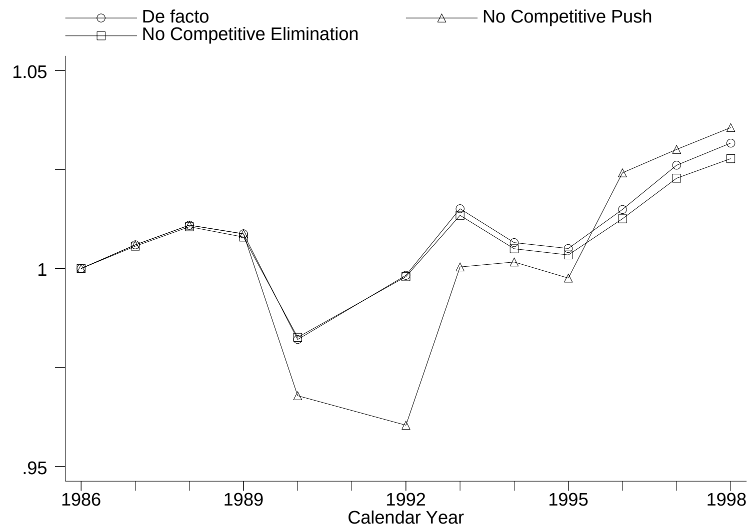
[b]Tariffs assumed to affect TFP change according to the estimate in table 3.4, column 1.

[c]Tariffs assumed to affect exit according to estimates in table 3.6, columns 3 and 6. In the counterfactual sample, an according share of exiting firms is randomly kept (with productivity at the level of *de facto* exit).

likely did not matter for productivity change because non-tariff barriers remained binding. In the following simulations, only the effect of tariffs is considered.

For the foreign input push (1), one needs to infer what share of foreign inputs after trade reform is due to lower tariffs. Firms did buy foreign equipment before trade liberalization. Lower tariffs will make the equipment cheaper, however, and boost demand. Simple regressions on the estimation sample show that market penetration rises 1.73 percentage points on average across all sectors when tariffs are reduced by 10 percentage points. In the machinery sector, however, market penetration almost triples (increases 289 percentage points) when tariffs are lowered by 10 percentage points. The estimate is likely upward biased because tariffs also catch the effect of the strongly changing non-tariff barriers. The estimate is still used. Even then, the simulated impact of foreign factors on productivity will be small.

The counterfactual share of foreign equipment (between 1990 and 1995) is calculated as $\hat{\kappa}^f_{i,t} = \kappa^f_{i,t} + .289 \cdot (\tau_t - \tau_{1988})$, where τ_t denotes nominal tariffs on investment goods in year t (measured on a scale from 0 to 1). This is a further favorable assumption for channel 1 since installed capital stocks would not respond as fast as equipment acquisitions. A similar calculation is applied to foreign intermediates (1996 through 1998). The results in table 3.7 corroborate that this channel is not important even under favorable assumptions. Except for a slight effect in 1990, productivity



Data: Simulated and *de facto* productivity in *Pesquisa Industrial Annual*.

Figure 3.4: Log TFP under three scenarios

would have evolved in about the same way had less foreign inputs been used.

To assess the competitive push (2), each individual firm's observed TFP is reduced by $\hat{\Delta} \ln \mathrm{TFP}_{i,t} = -2.249(\tau_t - \tau_{t-1})$ year over year between 1990 and 1998 (but not cumulatively). The coefficient estimate -2.249 for tariff levels is taken from table 3.4, column 1. Now, τ_t denotes nominal tariffs for products in that firm's sector. In 1990 only, TFP would have remained 1.4 percent lower had there not been an increase in foreign competition. Since tariffs were raised again after 1995, however, about .4 percent less TFP is observed in 1998 than would have been feasible had tariffs remained at their low.

For competitive elimination (3), exit among exporters and non-exporters is considered. Given the standardization chosen in the MNL model of section 3.5.3, the expected share of exits in year t can be expressed as $\mathbb{E}\left[n_{exit,t+1}/N_{active,t}|\tau_t\right] = \exp[\hat{\beta}_\tau(\tau_t - \tau_{1988})] \cdot P(\tau_{1988})$ where $\hat{\beta}_\tau$ is the coefficient estimate for tariffs (table 3.6, columns 3 and 6). Expected exit would be $\hat{\mathbb{E}}\left[\hat{n}_{exit,t+1}/N_{active,t}|\tau_{1988}\right] \equiv P(\tau_{1988})$ at 1988 tariffs. So, one can consider $1 - \exp[-\hat{\beta}_\tau(\tau_t - \tau_{1990})]$

an estimate for the relative share of exits $(n_{exit,t+1} - \hat{n}_{exit,t+1})/n_{exit,t+1}$ that is attributable to tariff reductions. To assess the counterfactual of frozen tariffs at the 1988 level, a share $1 - \exp[-\hat{\beta}_\tau(\tau_t - \tau_{1988})]$ was randomly drawn from the observed exiting firms and put back into the sample, duplicating their year t observation for $t+1$ and beyond, but allowing their productivity to rise along with average market productivity.

In the reported simulation, 432 observations of otherwise exiting firms are randomly added to the total sample of 65,621 observations. Surprisingly, simulated TFP in 1990 turns out to be slightly higher than *de facto* productivity on average. This may mean that trade reform induced high productivity firms to exit initially. Anecdotal evidence for the equipment sector, for instance, confirms this. Relatively advanced firms chose to exit in the early nineties since their products could often not compete with foreign goods, while domestic firms with low-quality and low-productivity products were favored. By 1998, the effect turns round. Had exiting firms stayed, productivity could have been up to .4 percent lower—close to a tenth of the *de facto* productivity change between 1990 and 1998.

Figure 3.4 depicts simulation results for the competitive push and competitive elimination. The competitive push (2) has a considerable and immediate impact on productivity. Competitive elimination (3) affects too few firms to have an immediate effect but has a sizeable impact over time. Qualitative evidence confirms this pattern. Amann (1999) investigates the Brazilian non-serial capital goods sector and argues that managers did not find foreign inputs with embodied technology a major source for innovation (p. 342). In other words, they did not expect much of a foreign input push (1). Amann (1999, p. 351) also states that managers restructured processes after 1989 but engaged in little efforts of their own to innovate products. This supports the importance of the competitive push (2). However some managers chose to obtain foreign designs to improve their products (Amann 1999, p. 342). That is a possibly important channel of knowledge flows unrelated to the flow of traded goods.

3.6 Conclusion

The Brazilian trade liberalization in the early nineties provides a natural setting to trace effects of international trade on productivity change. Brazil reduced tariffs and non-tariff barriers for imports but kept the treatment of exports largely unchanged. This gives the study a welcome focus. A sample of medium-sized to large Brazilian manufacturers is followed over the years from 1986 until 1998. Three channels through which trade reform affects productivity can be distinguished in the present data: (1) Easier access to foreign equipment and intermediates may allow for a *Foreign Input Push* at the firm level. (2) On the product market side, foreign imports constitute a *Competitive Push* on individual firms. Theory predicts that managers choose to innovate processes and remove slack under fiercer competition. (3) Competition in the product market may also induce more exits and cause a *Competitive Elimination* of inefficient firms.

The *Foreign Input Push* (channel 1) is found to be relatively unimportant. The efficiency difference between foreign and domestic inputs has only a minor impact on productivity. Foreign technology adoption takes time due to learning effects, factor complementarities and necessary production rearrangements. Estimates nevertheless support the view that foreign inputs create additional value, mostly in the medium term.

Trade liberalization induces high competitive pressure. First, it unleashes a *Competitive Push* on firms to raise their efficiency (channel 2). This proves to be a salient source of productivity change. Controlling for the endogeneity of foreign market penetration and tariffs, small changes in the tariff act are shown to induce impressive efficiency improvements among surviving firms. Second, when trade barriers fall, the *Competitive Elimination* of the least efficient firms (channel 3) strikes more fiercely. Estimates of turnover probabilities with multinomial logit confirm that both the likelihood of survival drops markedly when trade barriers fall and that low-efficiency firms go out of business more frequently. Counterfactual simulations indicate, however, that *Competitive Elimination* only exerts an impact on aggregate productivity over time. It stems from just a small share of firms. The simulations underscore the force of the *Competitive Push* on the other hand. This channel is a dominant source of productivity change among Brazilian manufacturers between 1990 and 1998.

Table 3.8: Foreign Competition and Productivity Change, Validity of Instruments

	Δ ln TFP[a]				ln TFP[a]			
	base IVs	add P_{Brazil}	add P_{USA}	add $P_{Argent.}$	base IVs	add P_{USA}	add $P_{Argent.}$	No IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tariff	-2.249	-1.909	-2.008	-1.604	-.609	-.516	.302	.184
	(.201)	(.119)	(.185)	(.115)	(.198)	(.145)	(.098)	(.021)
Mkt. Penetr.	27.989	22.595	24.316	22.903	10.652	9.695	2.438	1.468
	(2.163)	(1.485)	(1.981)	(1.596)	(1.904)	(1.312)	(1.076)	(.108)
κ^f	-.424	-.369	-.385	-.332	*-.043*	*-.033*	.055	*.033*
	(.053)	(.043)	(.049)	(.043)	(.034)	(.031)	(.027)	(.025)
μ^f	-.504	-.356	-.405	-.419	-.301	-.272	*-.07*	*-.004*
	(.074)	(.061)	(.068)	(.064)	(.063)	(.047)	(.043)	(.025)
ι(big L^{tot})[b]	.176	.168	.171	.182	.133	.129	.102	.087
	(.045)	(.041)	(.042)	(.041)	(.021)	(.02)	(.019)	(.018)
Demand[c]	-.548	-.492	-.509	-.466	.244	.261	.395	.392
	(.034)	(.024)	(.031)	(.024)	(.035)	(.026)	(.02)	(.01)
Cum. FDI[d]	.185	.148	.16	.138	.039	.031	-.03	-.029
	(.019)	(.014)	(.018)	(.014)	(.017)	(.012)	(.01)	(.005)
FDI flow[d]	-.14	-.121	-.127	-.113	*-.008*	*-.004*	.033	.033
	(.016)	(.013)	(.015)	(.013)	(.012)	(.01)	(.009)	(.007)
$\hat{\chi}^2$		98.5	-56.6	35.5		.634	192.4	
$Pr(\chi^2_{12} > \hat{\chi}^2)$		.0000	-	.0004		1.000	.0000	

[a] Regressors age, ι(medium L^{tot}), ι(medium cap.) and ι(big L^{tot}) not reported (see table 3.4).

[b] Labor force indicator: ($L^{tot}_{i,t} \geq 300$).

[c] Sector-wide sales in *PIA*, augmented by foreign market penetration.

[d] Billion USD per sector. Cumulated FDI is end-of-year stock of invested foreign capital.

Table 3.9: Further Multinomial Logit Estimates of Transition Probabilities

$\sigma_{i,t}$	Suspended Firm[a]			2S-IV: Non-Exporter[a]		
$\sigma_{i,t+1}$	Exp.	Dom.	Exit	Exp.	Susp.	Exit
	(1)	(2)	(3)	(4)	(5)	(6)
Tariff[b]	-18.728	-8.537	-4.264			
	(3.329)	(2.639)	(1.719)			
Real exch. rate (USD)[c]	11.012	8.441	4.661	-1.304	*-.359*	-3.851
	(2.551)	(1.993)	(2.118)	(.133)	(.433)	(.327)
ln TFP	.662	*-.032*	*-.218*	.076	-.281	-.236
	(.262)	(.223)	(.181)	(.019)	(.058)	(.037)
Age	*.024*	*-.017*	*-.002*	*.0001*	*-.003*	.007
	(.019)	(.018)	(.018)	(.001)	(.005)	(.003)
κ^f	6.621	*3.184*	*1.19*	1.215	*-.136*	*-.522*
	(3.338)	(3.495)	(2.91)	(.213)	(.846)	(.579)
μ^f				*-.001*	*-3.607*	*-.257*
				(.349)	(2.685)	(.716)
ι(big L^{tot})[d]	4.607	*1.503*	*.6*	1.158	-.882	-1.269
	(1.585)	(.869)	(.851)	(.181)	(.305)	(.192)
ι(big cap.)[e]	*-.101*	-2.067	*-.994*	1.242	*.006*	*-.114*
	(.951)	(.829)	(1.006)	(.116)	(.25)	(.172)
Sector demand	-.87	*-.1*	*-.1*	-.129	*-.034*	.193
	(.203)	(.123)	(.123)	(.013)	(.03)	(.02)
Cum. FDI[f]	*-.318*	*-.265*	*-.265*	*-.0008*	*-.096*	*.09*
	(.251)	(.34)	(.34)	(.028)	(.077)	(.051)
FDI flow[f]	*.898*	*.92*	*.92*	-.393	*.178*	*-.235*
	(1.263)	(.973)	(.973)	(.065)	(.178)	(.127)
Tariff[g]				*-.144*	-.994	-1.442
				(.14)	(.457)	(.388)
Mkt. Penetr.[g]				3.297	*2.158*	-14.636
				(1.43)	(4.044)	(2.524)
Obs.		146			27,391	
Pseudo R^2		0.302			0.665	
		$Pr(\chi^2_{36} > \hat{\chi}^2 = 137.68) = .0000$			$Pr(\chi^2_{42} > \hat{\chi}^2 = 22965.94) = .0000$	

[a] Regressors ι(medium L^{tot}) and ι(medium cap.) not reported (see table 3.6).

[b] Next year's tariff.

[c] Annual. Based on *IPA-OG* and US producer price index.

[d] Labor force indicator: ($L^{tot}_{i,t} \geq 300$).

[e] Capital stock indicator: $K_{i,t} + S_{i,t}$ in upper tercile of all firms in a year.

[f] Billion USD per sector. Cumulated FDI is end-of-year stock of invested foreign capital.

[g] Instrumented as in columns 2 and 3 of table 3.4.

Table 3.10: Production Function Estimates

	Olley-Pakes and further treatments							FE			
Nív.50	κ^f	k	s	μ^f	m	l^{wh}	l^{bl}	κ^f	k	μ^f	m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
04	*Non-metal mineral products* (OP: 2,074, FE: 2,775 observations)										
	-*.007*	.065	.061	-*.095*	.239	.167	.461	-*.005*	.067	-.625	.221
	(.071)	(.011)	(.014)	(.124)	(.016)	(.013)	(.019)	(.092)	(.013)	(.133)	(.01)
05	*Basic metal products* (OP: 500, FE: 648 observations)										
	-*.211*	*.021*	-*.0006*	*.847*	.243	.116	.491	-*.248*	*.021*	*.029*	.265
	(.179)	(.031)	(.04)	(.876)	(.042)	(.033)	(.047)	(.188)	(.034)	(.215)	(.02)
06	*Non-ferrous metal products* (OP: 341, FE: 689 observations)										
	-*.154*	.056	*.013*	*.152*	.315	.151	.461	-*.131*	.056	*.05*	.314
	(.197)	(.021)	(.017)	(.237)	(.032)	(.032)	(.039)	(.173)	(.024)	(.128)	(.016)
07	*Metal products* (OP: 1,890, FE: 2,503 observations)										
	.064	*.022*	.065	*.348*	*.227*	.17	.524	*.064*	*.022*	*.099*	.215
	(.055)	(.016)	(.017)	(.788)	(.263)	(.08)	(.092)	(.073)	(.013)	(.13)	(.009)
08	*Machinery and equipment* (OP: 1,742, FE: 2,291 observations)										
	-*.037*	.047	.077	*.154*	.222	.238	.439	-*.047*	.047	-*.03*	.229
	(.129)	(.02)	(.022)	(.098)	(.014)	(.018)	(.025)	(.12)	(.016)	(.088)	(.01)
10	*Electrical equipment* (OP: 1,552, FE: 2,042 observations)										
	-*.043*	*.019*	.106	*.075*	.188	.212	.384	-*.027*	*.019*	.326	.219
	(.158)	(.017)	(.014)	(.121)	(.012)	(.018)	(.028)	(.153)	(.018)	(.098)	(.011)

κ^f: share of foreign equipment, k: log of total equipment, s: log of other structures goods, μ^f: share of foreign intermediates, m: log of total intermediates, l^{wh}: log of number of white-collar workers, l^{bl}: log of number of blue-collar workers.

Standard errors: Estimates from 250 bootstraps.

Source: *Pesquisa Industrial Anual* 1986-98, deflated with *IPA-OG* as documented in appendix A.

Table 3.10: Production Functions, continued

	Olley-Pakes and further treatments							FE			
Nív.50	κ^f	k	s	μ^f	m	l^{wh}	l^{bl}	κ^f	k	μ^f	m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
11	*Electronic equipment* (OP: 934, FE: 1,207 observations)										
	.24	.061	.053	*.094*	.244	.27	.261	*.252*	.066	*.211*	.298
	(.11)	(.022)	(.025)	(.121)	(.039)	(.024)	(.039)	(.142)	(.022)	(.11)	(.013)
12	*Automobiles, trucks, and buses* (OP: 308, FE: 406 observations)										
	-*.272*	.113	*.055*	-*.032*	.156	.186	.588	-.567	.153	-*.132*	.201
	(.179)	(.043)	(.038)	(.406)	(.035)	(.043)	(.061)	(.228)	(.041)	(.242)	(.022)
13	*Other vehicles and parts* (OP: 1,249, FE: 1,621 observations)										
	.032	.089	*.043*	.237	.221	.178	.532	*.032*	.09	.465	.239
	(.075)	(.019)	(.022)	(.094)	(.017)	(.019)	(.027)	(.102)	(.018)	(.085)	(.01)
14	*Timber and furniture* (OP: 1,975, FE: 2,689 observations)										
	-.175	.116	.07	.689	.233	.163	.446	-.203	.124	-*.323*	.222
	(.068)	(.017)	(.014)	(.202)	(.015)	(.016)	(.027)	(.092)	(.015)	(.203)	(.01)
15	*Paper, pulp, and cardboard* (OP: 866, FE: 1,148 observations)										
	.146	.106	*.021*	*.521*	.267	.193	.386	*.118*	.113	-*.151*	.263
	(.194)	(.03)	(.02)	(.278)	(.028)	(.022)	(.03)	(.142)	(.017)	(.214)	(.013)
16	*Rubber products* (OP: 596, FE: 796 observations)										
	-*.387*	.063	.084	*.402*	.215	.163	.344	-*.377*	.064	-.491	.245
	(.274)	(.032)	(.019)	(.275)	(.036)	(.032)	(.044)	(.237)	(.029)	(.224)	(.021)

κ^f: share of foreign equipment, k: log of total equipment, s: log of other structures goods, μ^f: share of foreign intermediates, m: log of total intermediates, l^{wh}: log of number of white-collar workers, l^{bl}: log of number of blue-collar workers.

Standard errors: Estimates from 250 bootstraps.

Source: *Pesquisa Industrial Anual* 1986-98, deflated with *IPA-OG* as documented in appendix A.

Table 3.10: Production Functions, continued

Nív.50	Olley-Pakes and further treatments							FE			
	κ^f	k	s	μ^f	m	l^{wh}	l^{bl}	κ^f	k	μ^f	m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
17	*Non-petrochemical chemical elements* (OP: 1,071, FE: 1,412 observations)										
	-*.813*	.059	.093	*.103*	.413	.115	.226	-.823	.059	-.53	.412
	(.534)	(.027)	(.035)	(.146)	(.025)	(.014)	(.018)	(.309)	(.021)	(.153)	(.013)
18	*Basic petrochemical products* (OP: 790, FE: 1,025 observations)										
	-*.067*	.068	.099	*.052*	.186	.171	.3	-*.079*	.068	-.322	.135
	(.191)	(.02)	(.023)	(.133)	(.016)	(.017)	(.031)	(.169)	(.019)	(.11)	(.012)
19	*Chemical products* (OP: 1,111, FE: 1,422 observations)										
	.099	-*.006*	.047	*.053*	.156	.273	.458	*.067*	-*.006*	*.055*	.134
	(.155)	(.022)	(.019)	(.136)	(.018)	(.026)	(.029)	(.235)	(.019)	(.086)	(.013)
20	*Pharmaceutical products and perfumes* (OP: 1,201, FE: 1,570 observations)										
	-*.159*	.095	-*.014*	*.164*	.163	.167	.402	-*.258*	.093	-.613	.094
	(.113)	(.02)	(.019)	(.104)	(.015)	(.024)	(.029)	(.196)	(.023)	(.096)	(.013)
21	*Plastics* (OP: 1,492, FE: 1,964 observations)										
	-*.00009*	.033	.057	.606	.189	.223	.479	-*.00009*	.034	*.213*	.174
	(.095)	(.016)	(.014)	(.152)	(.019)	(.017)	(.024)	(.113)	(.014)	(.109)	(.01)
22	*Textiles* (OP: 2,483, FE: 3,225 observations)										
	.168	*.016*	.067	.258	.311	.15	.407	.168	*.016*	-.167	.297
	(.044)	(.014)	(.013)	(.067)	(.016)	(.015)	(.023)	(.059)	(.013)	(.076)	(.009)

κ^f: share of foreign equipment, k: log of total equipment, s: log of other structures goods, μ^f: share of foreign intermediates, m: log of total intermediates, l^{wh}: log of number of white-collar workers, l^{bl}: log of number of blue-collar workers.

Standard errors: Estimates from 250 bootstraps.

Source: *Pesquisa Industrial Anual* 1986-98, deflated with *IPA-OG* as documented in appendix A.

Table 3.10: Production Functions, continued

Nív.50	Olley-Pakes and further treatments							FE			
	κ^f	k	s	μ^f	m	l^{wh}	l^{bl}	κ^f	k	μ^f	m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
23	*Apparel* (OP: 1,559, FE: 2,091 observations)										
	.116	.128	.106	*.259*	.195	.215	.406	*.116*	.128	.388	.216
	(.072)	(.026)	(.02)	(.137)	(.016)	(.021)	(.025)	(.096)	(.018)	(.143)	(.01)
24	*Leather products and footwear* (OP: 1,371, FE: 1,745 observations)										
	.34	*.022*	*.023*	.995	.201	.127	.541	.255	*.024*	.521	.221
	(.086)	(.019)	(.016)	(.185)	(.023)	(.023)	(.034)	(.11)	(.018)	(.155)	(.014)
25	*Coffee products* (OP: 655, FE: 843 observations)										
	-.094	*.04*	*.047*	*.066*	.205	.294	.324	*-.1*	*.04*	*-.811*	.188
	(.142)	(.037)	(.035)	(3.043)	(.025)	(.036)	(.042)	(.292)	(.029)	(1.422)	(.016)
26	*Processed edible products* (OP: 2,233, FE: 2,927 observations)										
	-.322	.072	.048	.208	.246	.2	.332	-.318	.072	-.394	.224
	(.136)	(.018)	(.024)	(.082)	(.011)	(.018)	(.018)	(.151)	(.016)	(.086)	(.008)
27	*Meat and poultry* (OP: 896, FE: 1,194 observations)										
	.032	*.039*	*.01*	*.821*	.336	.152	.306	*-.079*	*.041*	*-.691*	.324
	(.179)	(.02)	(.025)	(.574)	(.04)	(.036)	(.062)	(.258)	(.024)	(.471)	(.016)
28	*Processed dairy products* (OP: 582, FE: 766 observations)										
	-.422	-.073	*.017*	*.793*	.276	.203	.332	*-.428*	-.073	*-.701*	.236
	(.202)	(.032)	(.028)	(1.225)	(.026)	(.03)	(.04)	(.237)	(.034)	(.696)	(.017)

κ^f: share of foreign equipment, k: log of total equipment, s: log of other structures goods, μ^f: share of foreign intermediates, m: log of total intermediates, l^{wh}: log of number of white-collar workers, l^{bl}: log of number of blue-collar workers.

Standard errors: Estimates from 250 bootstraps.

Source: *Pesquisa Industrial Anual* 1986-98, deflated with *IPA-OG* as documented in appendix A.

Table 3.10: Production Functions, continued

Nív.50	Olley-Pakes and further treatments							FE			
	κ^f	k	s	μ^f	m	l^{wh}	l^{bl}	κ^f	k	μ^f	m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
29	*Sugar* (OP: 336, FE: 757 observations)										
	-*.778*	.122	-*.02*	*1.669*	.304	.109	.158	-*.619*	.125	-*2.136*	.328
	(.434)	(.029)	(.027)	(7.152)	(.028)	(.02)	(.022)	(.339)	(.029)	(3.351)	(.017)
30	*Vegetable oil* (OP: 439, FE: 567 observations)										
	-.642	.075	.151	-*.344*	.356	.152	.257	-.702	.089	*.282*	.354
	(.26)	(.028)	(.037)	(.914)	(.035)	(.048)	(.044)	(.313)	(.028)	(.754)	(.021)
31	*Beverages and other food products* (OP: 2,468, FE: 3,236 observations)										
	-*.049*	.075	.04	*.29*	.213	.2	.396	-*.044*	.075	-.299	.188
	(.083)	(.013)	(.014)	(.151)	(.011)	(.015)	(.031)	(.132)	(.013)	(.134)	(.008)

κ^f: share of foreign equipment, k: log of total equipment, s: log of other structures goods, μ^f: share of foreign intermediates, m: log of total intermediates, l^{wh}: log of number of white-collar workers, l^{bl}: log of number of blue-collar workers.

Standard errors: Estimates from 250 bootstraps.

Source: *Pesquisa Industrial Anual* 1986-98, deflated with *IPA-OG* as documented in appendix A.

Part II

Information in Financial Markets

Chapter 4

Another look at information acquisition under fully revealing asset prices

A empirical puzzle arose from the Asian financial crisis and other crises. Why did the first attacks on East Asian currencies occur in mid 1997 and not before? As early as 1995 the Bank of International Settlements (BIS) had warned in reports that excessive domestic credit expansion in East Asian countries exposed investors to high risks. Even the International Monetary Fund (IMF) included subdued warnings for South Korea's financial sector in its annual report in 1995, and for Indonesia in 1996. So, information about financial weaknesses was publicly available early on, and it is hard to believe that the statements went unnoticed among key investors. Large investors may have had similar or better information even earlier.

Why would a crisis only occur years later? Suppose investors place high expectations on a gold mine. The price of stocks for this mine surges. One investor, however, finds out that the gold mine is empty. She has employed a superior research staff. Should she move today or silently go with the market? Her decision is not trivial. She faces a trade-off. If she sells some of her gold mine stocks, she makes her information known to other investors. This may cause a loss. The earlier the collapse occurs due to better information, the earlier the rise in stock prices will end, and the less gains this investor will reap. On the other hand, acting and thus revealing her information makes the timing of the collapse more predictable because other investors rationally infer her knowledge about the true fundamental. As a result of this trade-off, there may be incentives for the informed investor to go with the market for a substantial period and to turn around shortly before an anticipated collapse.

As simple as the example may seem, there are considerable obstacles to modelling the idea rigorously. In fact, the two chapters in this part will merely lay the grounds for a future model of information transmission in crises. A theory is needed that gives information an economic value. For this purpose, the present and following chapter consider the basic case of information acquisition.

4.1 A Problematic Paradox

One of the long living paradoxes in economics is Grossman's (1976) and Grossman and Stiglitz' (1980) assertion that no rational expectations equilibrium can exist if asset prices are fully revealing. The paradox proceeds in two steps: (i) No investor wants to buy information but rather extract everybody else's information from price if prices are fully revealing. But then (ii), if nobody acquires information, somebody has an incentive to do so. Grossman and Stiglitz (1980, *conjecture 6*) write: "In the limit, when there is no noise, prices convey all information, and there is no incentive to purchase information. Hence, the only possible equilibrium is one with no information. But if everyone is uninformed, it clearly pays some individual to become informed. Thus, there does not exist a competitive equilibrium." This conjecture and variants of it can be found in the literature ever since. Summarizing the belief succinctly, Barlevy and Veronesi (2000) remark: "Finally, as Grossman and Stiglitz point out, we need to prevent prices from being fully revealing; otherwise an equilibrium will fail to exist.” I will call this the ‘no equilibrium conjecture’ throughout part II of this dissertation.

Various ways around the paradox have evolved. Hellwig (1980) and Verrecchia (1982), for instance, add random shocks to the asset supply, and Kyle (1985) introduces liquidity traders and market makers. This way, they and many subsequent authors avoid the paradox by preventing the equilibrium price from becoming fully revealing. A second approach to overcome the paradox is to keep prices fully revealing but to force the agents, by equilibrium definition, to choose actions that are conditional on price in a different way. Milgrom (1981), for instance, shows that the paradox can be resolved in a model of bidding under such assumptions. Similarly, Dubey, Geanakoplos and Shubik (1987) and Jackson and Peck (1999) make agents submit demand functions in a Shapley-Shubik game so that a Nash equilibrium arises in which investors chose to acquire information. In a third approach, Hellwig (1982) and Routledge (1999) consider adaptive learning from past price so that, again, investors cannot condition on current price. Jackson (1991) pursues a fourth strategy. Instead of altering the assumptions on asset demand, he drops the price taking assumption and derives an equilibrium under fully revealing prices.

The present chapter addresses the no-equilibrium conjecture in a different way. There are no external random shocks to the price, there are no noise traders, and prices are fully revealing. Classical demand and price taking are retained as in any Walrasian rational expectations equilibrium (REE). This and the following chapter model the choice of information explicitly by using the theory of conjugate prior distributions (Raiffa and Schlaifer 1961). As opposed to Grossman’s (1976) and Grossman and Stiglitz’ (1980) initial models, and many of their followers, information choice is not imposed through an equilibrium definition that sorts investors into informed or uninformed ones. Investors have more than a binary choice. They can buy a (discrete) number of signals and thus have clear individual incentives for information acquisition. This approach turns information into a public good (or bad, under certain conditions) and makes the analysis amenable to standard economic tools. An REE does exist under fully revealing prices—both at Wall Street and in the market for information. It is unique when information is costly. Since markets for both assets and information (signals) clear, this REE is called a Rational Information Choice Equilibrium (RICE).

Intuitively, a RICE simply extends the common economic notion of an equilibrium to a market for information. I do not ask what bidding or clearing process leads to that equilibrium, whereas this may have been part of the motivation for Grossman and Stiglitz' (1980) statement of the no-equilibrium conjecture. Under their conjecture, however, information is like a pure public good. As Samuelson (1954) first argued, the competitive equilibrium for a public good is that amount of the public good along with that price for each unit of it where no agent wants to acquire any more or less. A RICE is an equilibrium exactly of this type, just in the context of rational expectations.

To make the framework concrete, this and the following chapter use CARA utility. Whereas the following chapter will employ only Gaussian random variables, the present chapter considers the more realistic case of a gamma distributed asset return. Poisson distributed signals make the model work smoothly. For information acquisition to occur in equilibrium, individuals need to be sufficiently risk averse, or asset supply relatively large, or the asset return sufficiently volatile. Under any of these conditions, information has a high value. So, the present model shows why part (i) of Grossman and Stiglitz' (1980) can fail: Even though others acquire information that will be publicly revealed, every single investor still has an individual incentive to buy information. The reason is that more information increases the precision of an investor's knowledge. So, the investor can make a more educated portfolio choice, and consumption tomorrow becomes less risky. Information behaves just like a standard public good, underprovided but provided. A benevolent social planner would implement more information.

However, information has a second and detrimental effect in financial markets when it becomes too common. From an investor's point of view, the value of the asset is given by its individually expected return less price. Commonly held information reduces this difference. As information gets mutually transmitted and extracted, individual beliefs and average market beliefs move closer to each other. Since asset price reflects average market opinion fully, it moves towards the individually expected return, diminishes the value of the asset and thus the value of information for each individual investor. This effect prevents information acquisition to occur in the companion model of chapter 5 with normally distributed random variables, and harms information acquisition in that model even when prices are not fully revealing. The negative effect can also prevail in the present model, and no information is acquired if investors are not very risk averse, markets are small, or returns are not very risky. In other words, it can also be the case that part (ii) of Grossman and Stiglitz' (1980) argument fails: There can be cases when investors do not want to acquire any information at all. Precisely because prices would be fully revealing, investors choose not to obtain any information so that nothing can get revealed in fact. This kind of equilibrium is informationally efficient: a benevolent social planner agrees that a public bad should not be accumulated.

Radner (1979) and Allen (1981) laid the grounds for REE under fully or partly revealing prices. These articles and a series of further contributions establish that a fully revealing rational expectations equilibrium at Wall Street generically exists for real assets (Jordan 1982, Pietra and Siconolfi 1998, Citanna and Villanacci 2000a) but not necessarily for nominal assets (Rahi 1995).[1] However, these articles stop short of investigating the resulting incentives for investors to acquire information. The findings in the present chapter are reassuring: Investors do want to buy information even if asset prices are fully revealing. As in those articles, investors are assumed to be price takers in the present model. An extension to imperfect competition is left for future investigation.

To reflect the practical process of asset trading more closely than an REE can, Kyle (1985) and Back (1992), Admati and Pfleiderer (1988), and Easley and O'Hara (1992), to name just a few, add market makers (besides liquidity traders). Market makers call a price that equals the expected asset value, given their information from observing order flow, and prices become fully revealing only when trading stops in the final period. In such a setting, Foster and Viswanathan (1993) and Holden and Subrahmanyam (1996) give investors a choice of information. Their equilibrium concept

[1] Wang (1993), Einy, Moreno and Shitovitz (2000), Citanna and Villanacci (2000b) and many others investigate the informational properties of REE—that is, how partly or fully revealing prices aggregate information that became available.

resembles the one of Grossman and Stiglitz (1980) in that investors have only a binary choice of becoming informed or remaining uninformed. While disregarding the market making process and returning to REE for tractability, the present model and its companion version in chapter 5 give investors the choice of a discrete number of signals.

Of course, the models of this and the following chapter also relate to the huge body of alternative approaches to investor behavior. Banerjee (1992), Benabou and Laroque (1992), Caplin and Leahy (1994) and Avery and Zemsky (1998), for instance, rationalize herding behavior in financial markets. Incentives for information acquisition have also been analyzed in the context of social learning and experimentation (Burguet and Vives 2000, Moscarini and Smith 2001, Bergemann and Välimäki 2002). Calvo and Mendoza (2000) and Popper and Montgomery (2001) investigate the occurrence of crises in the context of informational asymmetries. However, none of these models internalizes the optimal choice of information in a rational expectations equilibrium.[2] That is the focus of the framework in this and the next chapter.

The following section 4.2 builds a model of an investor's portfolio choice. A unique financial market equilibrium results under CARA utility and a gamma distributed asset return. Every investor can buy Poisson distributed signals prior to the portfolio choice. Section 4.3 analyzes this information choice and shows that a unique equilibrium exists in the market for signals, too. Under conditions spelled out, the equilibrium entails a positive amount of information. However, it is informationally inefficient under these conditions as shown in section 4.4. Section 4.5 concludes.

4.2 The Model

There are only two periods, today and tomorrow, and there are only two assets: One riskless bond and one risky stock. When Wall Street opens at 10am today, investors can choose their portfolio. The assets will yield a payoff tomorrow once and for all. All investors hold prior beliefs

[2] Rational expectations equilibria that internalize information transmission in product markets are analyzed in a complementary line of research for oligopoly firms (Clarke 1983, Gal-Or 1985, Raith 1996).

about the distribution of the risky asset return. Newsstands in New York, where all investors happen to live, open at 9am today. How many different private detectives should an investor hire? Each investor knows that she will base her portfolio decision, to be taken at 10am today, on the information that she is about to get out of her private detectives. She also knows the statistical distribution of the information that she can expect from the private detective, which is more informative than her own prior beliefs. But, of course, she does not know what exactly is going to be written in the private detective's report when she takes her decision on information acquisition. Otherwise she would not need to acquire the information.

The asset price at 10am will contain information. The reason is that each investor takes her portfolio decision given her information, and the Walrasian auctioneer at Wall Street makes the markets clear by calling an equilibrium price. So, each investor knows that the equilibrium asset price at 10am will reflect the information that all other investors will have received since those others base their asset demand on their respective information, too. In the most extreme case, the asset price at 10am will *fully reveal* everybody's information. This is the case of concern in the present chapter. The price will allow every investor to extract a sufficient statistic that summarizes all investors' posterior beliefs.

Let investor i maximize expected utility

$$U^i = \mathbb{E}\left[u(C_0^i) + \delta u(C_1^i)\,\middle|\,\mathcal{F}^i\right] \tag{4.1}$$

with respect to consumption C_0^i today and C_1^i tomorrow, and a portfolio choice. $\mathcal{F}^i$ denotes the information set available to investor i at the time of her portfolio choice. For ease of notation, abbreviate investor i's conditional expectations with $\mathbb{E}^i[\cdot] \equiv \mathbb{E}\left[\cdot\,\middle|\mathcal{F}^i\right]$ when they are based on posterior information, and with $\mathbb{E}^i_{prior}[\cdot] \equiv \mathbb{E}\left[\cdot\,\middle|\mathcal{F}^i_{prior}\right]$ for prior information. Posterior expectations will ultimately coincide for all investors under fully revealing prices, but it is instructive to keep them different for the derivation.

Both assets are perfectly divisible. The riskless bond pays a fixed interest rate r tomorrow

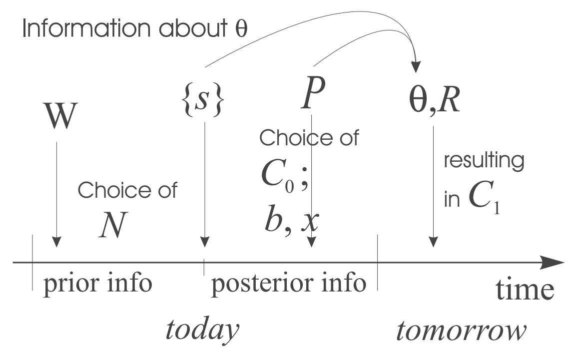


Figure 4.1: Timing of information revelation and decisions

so that the interest factor is $R \equiv 1+r \in [0, \infty)$. The risky asset pays a gross dividend θ tomorrow. Then, the intertemporal budget constraint of investor i becomes

$$b^i + Px^i = W_0^i - C_0^i - cN^i \tag{4.2}$$

today, and

$$C_1^i = Rb^i + \theta x^i \tag{4.3}$$

will be available for consumption tomorrow. The investor is endowed with initial wealth W_0^i, and chooses her consumption C_0^i and C_1^i, her holdings of the riskless bond b^i, her holdings of the risky stock x^i, and how much information N^i she wants. N^i denotes the number of signals (private detectives) that investor i chooses to contract. Each signal has a cost of c and conveys unbiased information about the return θ of the risky asset tomorrow.

The timing of decisions is illustrated in figure 4.1. First, investors have to choose the number of signals (private detectives) N^i. To do so, they maximize *ex ante* utility based on their prior information. Investors then receive the realizations $\{s_1^i, ..., s_{N^i}^i\}$ of these N^i signals (they get to know the content of the private detective's report), and update their beliefs. Then they choose consumption and savings, and decide how many risky assets to hold. At this stage, they maximize *ex* *ante* utility based on their posterior information. The Walrasian auctioneer in the financial market sets the price P for the risky asset such that the stock market clears. The world bond market clears by assumption, given the world interest rate r.

The present two-period model of utility maximization is similar to a model of terminal wealth maximization as in Grossman and Stiglitz (1980). However, beyond terminal wealth maximization, investors can also choose consumption today. This allows for an analysis of the value of information when the interest rate changes. In equilibrium, the relative price P of risky assets aggregates the information and can, under certain conditions, fully reveal the information of everybody in the market. All investors know the structure of the economy, which in turn is common knowledge. Beyond this fundamental assumption, the following conditions will be shown to be necessary and sufficient for prices to become fully revealing.

Assumption 4.1 (Common risk aversion) *Investors are risk averse and share a common and certain degree of risk aversion.*

Assumption 4.2 (Common priors) *Investors hold the same prior beliefs about the joint distribution of the risky asset return and the signals.*

Assumption 4.3 (Finite risk) *The prior variance of the risky asset return is strictly positive and finite.*

Assumption 4.4 (No borrowing constraint) *Investors can carry out unlimited short sales.*

The following assumptions are made to address the no-equilibrium conjecture in its pure form.

Assumption 4.5 (Known market size) *The total supply of the risky asset $\bar{x}$ and the total number of investors I are certain and known.*

Assumption 4.6 (Exogenous asset and signal supply) *Supply of the risky asset $\bar{x}$ is finite and strictly positive. Supply of the riskless asset and supply of the signals are perfectly elastic.*

Assumption 4.7 (Unique information) *All signals* $\{S_1^i, ..., S_{N^i}^i\}_{i=1}^I$ *are conditionally independent given the realization of the asset return,* $S_n^i|\theta \overset{i.i.d.}{\sim} f(s_n^i\,|\theta)$.

Assumption 4.8 (Equal precision of signals) *All signals have equal precision given the realization of the asset return.*

Assumption 4.9 (Price taking) *Investors are price takers in all markets.*

In addition to those, the following assumptions are made to make the model tractable and interesting.

Assumption 4.10 (CARA) *Investors have CARA utility with* $u(C) = -e^{-AC}$.

Assumption 4.11 (Distributions) *The risky asset return is gamma distributed, and signals are Poisson distributed.*

The particular choice of Poisson distributed signals implies that they are equally precise (assumption 4.8), given the asset return they inform about. The choice of any conjugate prior distribution such as in assumption 4.11 implies that signals are conditionally independent (assumption 4.7).

4.2.1 Investors' beliefs

Investors can have different information about the two parameters α^i, β^i of the risky asset's gamma distribution.[3] Hence, from an individual investor's point of view, the risky asset return is distributed $\theta \sim \mathcal{G}(\alpha^i, \beta^i)$ so that its p.d.f. is

$$\pi\left(\theta\left|\alpha^i, \beta^i\right.\right) = \begin{cases} \frac{(\beta^i)^{\alpha^i}}{\Gamma(\alpha^i)}\,\theta^{\alpha^i - 1}\,e^{-\beta^i\cdot\theta} & \text{for } \theta > 0 \\ 0 & \text{for } \theta \leq 0 \end{cases},$$

written in the style of Bayesian statistics.The two parameters α^i and β^i must be positive. A gamma distributed asset return has several nice features. Above all, the return is restricted to positive

[3]Earlier models in finance that employ the gamma distribution include Davis (1993), Knight, Satchell and Tran (1995), and Browne (1995) who shows that returns from particular trading strategies approach a gamma distribution.

realizations which makes the gamma distribution a more realistic distribution of the asset return than the normal distribution.

The expected value of a gamma distributed return θ is α^i/β^i, and its variance $\alpha^i/(\beta^i)^2$. The mean-variance ratio will play a key role in particular: $\mathbb{E}^i[\theta]/\mathbb{V}^i(\theta) = \beta^i$. The gamma distribution has the convenient property that a sum of M independently gamma distributed random variables $X_1 + ... + X_M$ with parameters $\alpha_1, ..., \alpha_M$ and β is itself a gamma distributed random variable with parameters $\alpha_1 + ... + \alpha_M$ and β (DeGroot 1989, Ch. 5.9). So, the following results generalize to the case where investors have a choice among M risky assets, each with a gamma distributed return θ_m and parameters α^i_m and β^i $(k = 1, ..., M)$. Section B.2 in appendix B (p. 261) lists properties of the gamma distribution that are useful in the following derivations.

A system of signals can be such that the posterior distribution of θ is again a gamma distribution. Statisticians call a distribution a *conjugate prior distribution* to the signals' distribution if this is satisfied. The gamma distribution is a conjugate prior to the Poisson distribution, to the normal distribution, and to itself, for instance.[4] Let signals be Poisson distributed $S^i_n|\theta \overset{i.i.d.}{\sim} \mathcal{P}(\theta)$ so that their p.d.f. is

$$f\left(s^i_n\,|\theta\right) = \begin{cases} e^{-\theta}\,\theta^{s^i_n}\frac{1}{s^i_n!} & \text{for } s^i_n > 0 \\ 0 & \text{for } s^i_n \leq 0 \end{cases}.$$

Both the mean and the variance of a Poisson distributed variable S^i with parameter θ are equal to θ. This can be a limitation for some modelling purposes. Not only the mean but also the variance of the signals is determined by the realization of the Poisson parameter θ. Thus, the precision of a signal $\mathbb{E}^i_{prior}\left[\mathbb{V}^i\left(s^i|\theta\right)\right]^{-1} = \mathbb{E}^i_{prior}[\theta]^{-1} = \beta^i_{prior}/\alpha^i_{prior}$ depends solely on individual priors and cannot be altered. As a consequence, signals cannot be modelled as more or less informative. However, this property of the gamma-Poisson conjugate pair does not impede the derivation of the main results in the present chapter. In fact, it simplifies the analysis. The sum of N^i conditionally independent Poisson signals is itself Poisson distributed with mean and variance $N^i\theta$ (section B.2

[4]See section B.1 in appendix B for general remarks on conjugate prior distributions.

in appendix B).

In day-to-day language, the term information can take two meanings. When we say we acquire information, we mean that we buy a signal, not knowing its realization. For if we knew the realization, we would not pay for it. However, when we say we have received information, we usually mean that we have learnt a signal's realization. Throughout this part II of the dissertation, I will continue to use the term information in both meanings to make the language less cumbersome. But I will keep the distinction between a signal and its realization clear. There is one key feature of signals and private detectives in the present model. Every investor receives a conditionally independent signal S_n^i under conjugate prior distributions. So, there is only one copy S_n^i of every private detective's report in the present model economy. In other words, a private detective can only be hired by one investor at a time. The following chapter 5 will vary this assumption and draw and analogy between signals and newspapers.

Suppose all investors have the same priors about the distribution of θ so that $\alpha^i_{prior} = \bar{\alpha}$ and $\beta^i_{prior} = \bar{\beta}$. Then, Poisson distributed signals result in an intuitive relationship between the prior and the posterior distribution of the asset return.

Fact 4.1 *Suppose the prior distribution of θ is a gamma distribution with parameters $\bar{\alpha} > 0$ and $\bar{\beta} > 0$. Signals $S_1^i, ..., S_{N^i}^i$ are independently drawn from a Poisson distribution with the realization of θ as parameter. Then the posterior distribution of θ, given realizations $s_1^i, ..., s_{N^i}^i$ of the signals, is a gamma distribution with parameters $\alpha^i = \bar{\alpha} + \sum_{n=1}^{N^i} s_n^i$ and $\beta^i = \bar{\beta} + N^i$.*

Proof. See DeGroot (1989, Ch. 6.3). ■

Fact 4.1 has an immediate implication for the *ex ante* variance of the asset return. For risk averse investors to have an incentive and acquire information at all, it is important that the *ex ante* variance is falling in the number of signals N^i. Indeed,

$$\frac{\partial}{\partial N^i}\mathbb{E}^i_{prior}\left[\mathbb{V}\left[\theta\left|\alpha^i, \beta^i\right.\right]\right] = \frac{\partial}{\partial N^i}\left(\frac{\bar{\alpha} + \frac{\bar{\alpha}}{\bar{\beta}}N^i}{(\bar{\beta} + N^i)^2}\right) = -\frac{\bar{\alpha} + \frac{\bar{\alpha}}{\bar{\beta}}N^i}{(\bar{\beta} + N^i)^3} < 0$$

(by fact B.5 in appendix B, p. 262). This is good news for risk averse individuals: Investors can lower the *ex ante* variance of the risky asset return so that they will be able to make a more educated portfolio choice at 10am. Since investors anticipate this improved portfolio choice at 9am, they consider information acquisition a means of reducing the *ex ante* variance of tomorrow' consumption.

4.2.2 The financial market equilibrium

Let's restrict attention to the equilibrium at Wall Street for now and disregard the market for private detectives. Suppose investors $i = 1, ..., I$ have received a discrete number of signals $N^i \geq 0$ each. It's 10am now, and they choose their portfolios (b^i, x^i) given their respective posterior information sets $\mathcal{F}^i$. In an REE, investors do not only consider the information that they get out of their own signals. They simultaneously extract information from price so that $\mathcal{F}^i = \{\sum_{n=1}^{N^i} s_n^i, P\}$. Since P and $\sum_{n=1}^{N^i} s_n^i$ are correlated in equilibrium, the posterior distribution of the asset return, based on this information set, is complicated. If price P is fully revealing, however, the information sets of all investors will coincide: $\mathcal{F}^i = \mathcal{F} = \{\sum_{k=1}^{I} \sum_{n=1}^{N^k} s_n^k\}$ for all i. This will give the rational beliefs in REE a simple linear form analogous to fact 4.1.

For CARA utility and a degree of absolute risk aversion A, and since

$$C_1^i - C_0^i = (1+R)b^i + (\theta + P)x^i - W_0^i + cN^i$$

by (4.2) and (4.3), the first order conditions for the optimal portfolio $(b^{i,*}, x^{i,*})$ become

$$\begin{aligned} \frac{1}{\delta} &= R\,\mathbb{E}^i\left[e^{-A(C_1^{i,*} - C_0^{i,*})}\right] = R \cdot H^i\,\mathbb{E}^i\left[e^{-Ax^{i,*}\cdot\theta}\right] && (4.4) \\ \frac{P}{\delta} &= \mathbb{E}^i\left[\theta\, e^{-A(C_1^{i,*} - C_0^{i,*})}\right] = H^i\,\mathbb{E}^i\left[\theta \cdot e^{-Ax^{i,*}\cdot\theta}\right] && (4.5) \end{aligned}$$

where $H^{i,*} \equiv \exp\left(-A\left[(1+R)b^{i,*} + Px^{i,*} - W_0^i + cN^i\right]\right)$ is certain since price is certain in a Walrasian REE. These first-order conditions hold irrespective of the distribution of the asset return.

Dividing (4.5) by (4.4) and using facts B.2 and B.3 (section B.2 in appendix B, p. 261), yields demand for the risky asset

$$x^{i,*} = \frac{\beta^i}{A} \frac{\mathbb{E}^i [\theta] - RP}{RP}. \tag{4.6}$$

for a gamma distributed asset return. By fact B.1 (appendix B, p. 261), the expected value of the return is $\mathbb{E}^i [\theta] = \alpha^i / \beta^i$. As it should be, demand for the risky asset is rising whenever its price falls or the riskless asset's return falls; demand is the higher the less risk averse investors become (lower A) or the higher the expected mean-variance ratio β^i of the asset is. Investors go short in the risky asset whenever their return expectations fall short of opportunity cost, $\mathbb{E}^i [\theta] < RP$, and go long otherwise. Due to the assumption of constant absolute risk aversion, demand for the risky asset is independent of the investor's wealth.

The term $\mathbb{E}^i [\theta - RP]$ is an individual investor i's *expected excess return* over opportunity cost. Risk averse investors demand this premium. This term and the closely related ratio $\mathbb{E}^i [\theta - RP] / RP$ have important informational properties that will affect the incentives for information acquisition in section 4.3. Define

$$\xi^i \equiv \frac{\mathbb{E}^i [\theta] - RP}{RP}, \tag{4.7}$$

the *expected relative excess return*. When price is informative, the expectations of the average market participant are closely reflected by asset price P and thus RP. Price is high when average market expectations are good, and vice versa. This can create a negative incentive for acquiring information. As private information becomes at least partly known to investors when prices are informative, RP moves closer to $\mathbb{E}^i [\theta]$ so that the risky asset loses value for an individual investor, and consequently information does.

Taking signal choice in the private detective market as given for now, a (partial) REE at Wall Street can be defined as a price P and consistent posterior beliefs based on $\mathcal{F}^i = \{\sum_{n=1}^{N^i} s_n^i, RP\}$ such that the portfolio $(b^{i,*}, x^{i,*})$ is optimal and asset markets clear, $\sum_{k=1}^{I} x^{k,*} = \bar{x}$.

Proposition 4.1 *Under assumptions 4.1 through 4.11, the unique rational expectations equilibrium*

(REE) in the asset market is

$$\alpha^i = \bar{\alpha} + \sum_{k=1}^{I} \sum_{n=1}^{N^k} s_n^k \equiv \alpha, \tag{4.8}$$

$$\beta^i = \bar{\beta} + \sum_{k=1}^{I} N^k \equiv \beta, \tag{4.9}$$

$$RP = \frac{I\alpha}{A\bar{x} + I\beta}. \tag{4.10}$$

Proof. By (4.6) and for beliefs (4.8) and (4.9), $x^* = \alpha/(A\,RP) - \beta/A$ for all i. So, market clearing $Ix^* = \bar{x}$ implies (4.10).

Uniqueness of beliefs (4.8) and (4.9) follows by construction. Price is fully revealing under assumptions 4.1 through 4.5 since there is no random component to price other than signal realizations. So, by (4.6) and market clearing, RP can always be written as $RP = T_0 + T_1(\sum_{k=1}^{I} \sum_{n=1}^{N^k} s_n^k)$ for an appropriate choice of constants $T_0, T_1 > 0$ because risk aversion A is common to all investors. But then, every investor i can infer $\sum_{k \neq i} \sum_{n=1}^{N^k} s_n^k = (RP - T_0)/T_1 - \sum_{n=1}^{N^i} s_n^i$ from her knowledge of own signal realizations. Since the random variables $\sum_{k \neq i} \sum_{n=1}^{N^k} s_n^k$ and $\sum_{n=1}^{N^i} s_n^i$ are Poisson distributed by fact B.5 (appendix B, p. 262) and conditionally independent given θ, a rational investor must apply Bayesian updating following fact 4.1 (appendix B, p. 111). Hence, $\alpha^i = \bar{\alpha} + \sum_{n=1}^{N^i} s_n^i + \sum_{k \neq i}^{I} \sum_{n=1}^{N^k} s_n^k$ and $\beta^i = \bar{\beta} + N^i + \sum_{k \neq i}^{I} N^k$.

Finally, no less than $\sum_{k=1}^{I} \sum_{n=1}^{N^k} s_n^k$ signals can get revealed in REE. Suppose one signal s_n^i is received by some investor i but does not enter price. Then, investor i cannot have based demand x^i on that signal since market clearing $\sum_{k=1}^{I} x^{k,*} = \bar{x}$ would have transmitted s_n^i to price. However, if α^i does not include s_n^i, Bayesian updating following fact 4.1 is violated, which is ruled out in an REE. ∎

So, the equilibrium price P fully reveals the *aggregate* information of all market participants. Knowing everything else, investors can rationally infer from RP what average market expectations are. Formally, aggregate information is the total of all signals received: $\sum_{i=1}^{I} \sum_{n=1}^{N^i} s_n^i$—a sufficient statistic for all moments of θ given $\sum_{i=1}^{I} N^i$. If the risky asset were not supplied autonomously so that $\bar{x} = 0$, then $\mathbb{E}^i[\theta] = RP$ under fully revealing prices. As a consequence, no risky assets would be demanded in equilibrium. Assumption 4.6 ($\bar{x} > 0$) prevents this. In general, the equilibrium price is fully revealing only if the assumptions are satisfied that have been made along the way. The following corollary restates them for the present context.

Corollary 4.1.1 *Suppose utility is CARA, the asset return is gamma distributed and signals are Poisson. Then equilibrium price P fully reveals all market participants' information* $\sum_{i=1}^{I}\sum_{n=1}^{N^i} s_n^i$ *if and only if*

- *assumptions 4.1 through 4.5 are satisfied, and*
- *the total number of all other investors' signals* $\sum_{k=1}^{I} N^k$ *is known to each investor i at the time of the portfolio choice.*

Proof. Sufficiency was established in proposition 4.1. Necessity of assumptions 4.1 through 4.3 and 4.5 follows by inspection of the general solution for market price given individual beliefs $\alpha^i = \bar{\alpha} + \sum_{n=1}^{N^i} s_n^i$ and $\beta^i = \bar{\beta} + N^i$:

$$RP = \frac{\frac{1}{I}\sum_{i=1}^{I}\frac{\alpha^i}{A^i}}{\frac{\bar{x}}{I} + \sum_{i=1}^{I}\frac{\beta^i}{A^i}} = \frac{\bar{\alpha}\left(\frac{1}{I}\sum_{i=1}^{I}\frac{1}{A^i}\right) + \frac{1}{I}\sum_{i=1}^{I}\frac{1}{A^i}\sum_{n=1}^{N^i} s_n^i}{\frac{\bar{x}}{I} + \bar{\beta}\left(\frac{1}{I}\sum_{i=1}^{I}\frac{1}{A^i}\right) + \frac{1}{I}\sum_{i=1}^{I}\frac{1}{A^i}N^i}.$$

If $\sum_{k=1}^{I} N^k$ were unknown to investor i, she would not be able to extract the sufficient statistic $\sum_{k=1}^{I}\sum_{n=1}^{N^k} s_n^k$ from price.

For necessity of assumption 4.4, consider the case in which some investors cannot go short in the risky asset due to a borrowing constraint. Then another investor will not know whether the equilibrium price is low because many relatively poor investors received bad signals and hit their borrowing constraint or whether only a few relatively wealthy investors received extremely bad signals. As a consequence, uncertainty remains and the price cannot be fully revealing. ■

Investors need to know how many reports of private detectives were read in total, otherwise prices at Wall Street will not be fully revealing. The definition of a rational information choice equilibrium (RICE) will assure that the total number of reports from private detective becomes common knowledge. As a consequence of CARA utility, it does not matter for the informativeness of price whether an investor knows everybody else's wealth. Note that price taking behavior is not a necessary condition for asset prices to be fully revealing (see Jackson (1991)).

4.3 Information Equilibrium

How much information do investors buy in equilibrium? The present framework provides a clear criterion: Each investor will purchase information until the cost of an additional signal exceeds its *ex ante* utility benefit, given prior beliefs and everybody else's signal choice. When prices are fully revealing, information is a public commodity. Like in any problem of public good provision, only the cost of providing the public good can be allocated while the good itself is available to everybody. Similarly, the allocation of signals at 9am is well defined but asset price at 10am will make signal realizations publicly available.

Definition 4.1 (Rational Information Choice Equilibrium) *A rational information choice equilibrium (RICE) is an allocation of $x^{i,*}$ risky assets, $b^{i,*}$ riskless bonds, and $N^{i,*}$ signals to investors $i = 1, ..., I$ and an asset price P along with consistent beliefs such that*

- *the portfolio $(x^{i,*}, b^{i,*})$ is optimal given RP and investors' posterior beliefs for $i = 1, ..., I$,*
- *the market for the risky asset clears, $\sum_{i=1}^{I} x^{i,*} = \bar{x}$, and*
- *the choice of signals $N^{i,*}$ is optimal for investors $i = 1, ..., I$ given the sum of all other investors' signal choices $\sum_{k \neq i} N^{k,*}$.*

An equilibrium with information choice must have two stages since signal realizations cannot be known at the time of signal acquisition. On the first stage, investors choose the number of signals given the choice of all other investors, and a Bayesian Nash equilibrium results. Since information is a public commodity under fully revealing price, each investor needs to take everybody else's choice into account. On the second stage, investors are price takers here. That is, they consider the impact of their demand on equilibrium price as negligible, and a competitive Walrasian REE results given the Bayesian Nash equilibrium on the first stage. One could call this a hybrid equilibrium since investors behave as price takers on the second change but rationally anticipate on the first stage that each investor's small contribution to price makes it fully revealing in the aggregate (investors are also price takers with respect to signal cost). Hellwig (1980) called this behavioral contrast in the assumptions 'schizophrenic'. By modelling games in which entire demand functions are submitted to the auctioneer, Dubey et al. (1987), Kyle (1989), Jackson (1991) and Jackson and Peck (1999), for instance, address this issue. The objective of the present chapter is different. It is to show that none of the classical assumptions needs to be dropped, and information will still be acquired under a fully revealing price.

In general, investors must act on information in equilibrium (at least for a nontrivial set of signal realizations). If they knew they would not respond to information, the anticipated equilibrium at Wall Street would be exactly the same with or without their signal acquisition. So, signals would not have any benefit but cause a cost, and the investor would not buy any. Hence, even when investors choose their portfolio strategically they must anticipate responding to signal realizations at least partially.[5] In the above-defined REE, they respond fully to signal realizations. In this sense, price-taking behavior is the benchmark case. Investors neglect the small but existent impact of their individual demand on price when they choose the portfolio. Their many little contributions to aggregate demand still result in a price that reveals the overall information behind all these individually negligible responses, and investors rationally use this information.

[5] Jackson (1991) derives an REE under an exponential distribution of the asset return, a special case of the gamma distribution, with risk neutral agents who strategically adjust asset demand. Jackson shows that the resulting price is fully revealing despite investors' strategic behavior.

Asset demand and price in REE were derived in the preceding section 4.2.2. The equilibrium amount of information remains to be determined. For CARA, expected indirect utility is

$$U^{i,*} \quad = \quad -\tfrac{1+R}{R}(\delta R)^{\frac{1}{1+R}}\, e^{-A\frac{R}{1+R}(W_0^i - cN^i)}\, \mathbb{E}^i\left[e^{-Ax^i(\theta-RP)}\right]^{\frac{1}{1+R}} \tag{4.11}$$

at 10am, irrespective of the distribution of the risky asset return (see section B.3 in appendix B, p. 262). It becomes

$$U^{i,*} = -\tfrac{1+R}{R}(\delta R)^{\frac{1}{1+R}}\, e^{-A\frac{R}{1+R}(W_0^i - cN^i)} \left(\frac{e^{\beta^i \mathbb{E}^i[\theta-RP]}}{\mathbb{E}^i\left[\frac{\theta}{RP}\right]^{\alpha^i}}\right)^{\frac{1}{1+R}} \tag{4.12}$$

for gamma distributed returns in particular (section B.3 in appendix B, p. 263). At the time of the signal choice, signal realizations have not been transmitted to investors yet. Therefore, investors base their expectations of $U^{i,*}$ on their priors about the distribution of θ and their priors about the distribution of the signals $\{S_1^k, ..., S_{N^k}^k\}_{k=1}^I$ at 9am. So, when investors choose the number of signals, they maximize *ex ante* utility $\mathbb{E}^i_{prior}\left[U^{i,*}\right]$ with respect to the number of signals. (To be precise, they maximize expected indirect utility based on prior beliefs.) *Ex ante* utility takes the form

$$\begin{aligned} \mathbb{E}^i_{prior}\left[U^{i,*}\right] \quad &= \quad -\tfrac{1+R}{R}(\delta R)^{\frac{1}{1+R}} \cdot e^{-A\frac{R}{1+R}(W_0^i - cN^i)} \\ &\quad \cdot \mathbb{E}^i_{prior}\left[\left(\frac{\alpha^i}{\sum_i \alpha^i}\frac{A\bar{x}+\sum_i \beta^i}{\beta^i}\right)^{-\frac{\alpha^i}{1+R}} e^{\frac{1}{1+R}\left(\alpha^i - \frac{\beta^i}{A\bar{x}+\sum_i \beta^i}\sum_i \alpha^i\right)}\right] \end{aligned} \tag{4.13}$$

since $\beta^i\mathbb{E}^i\left[\theta - RP\right] = \alpha^i - \beta^i RP$ and $\mathbb{E}^i\left[\theta/RP\right] = \alpha^i/(\beta^i RP)$ by fact B.1 (section B.2 in appendix B, p. 261), while $RP = (\sum_i \alpha^i)/(A\bar{x} + \sum_i \beta^i)$ by (4.10).

At the time of information acquisition (9am), the anticipated posterior parameter (4.9) $\beta^i = \beta = \bar{\beta} + N^i + \sum_{k\neq i}^I N^k$ is already known to every investor i since she knows how many signals she purchased herself and the aggregate choice of everybody else. The posterior parameter α^i (4.8), on the other hand, is uncertain *ex ante*. It depends on the signal realizations, which are to arrive. So, α^i is Poisson distributed conditional on the signals' mean θ (the Poisson parameter). In addition, the Poisson parameter is gamma distributed, with prior parameters $\bar{\alpha}$ and $\bar{\beta}$. If prices are fully revealing, then the Bayesian updated parameter α^i of the posterior distribution is the same for all investors and equal to $\alpha^i = \alpha = \bar{\alpha} + \sum_{k=1}^I \sum_{n=1}^{N^k} s_n^k$. Hence, iff prices are fully revealing, the last term in (4.13) simplifies to

$$\begin{aligned}
&\mathbb{E}^i_{prior}\left[\left(\frac{\alpha}{I\alpha}\frac{A\bar{x}+I\beta}{\beta}\right)^{-\frac{\alpha}{1+R}} e^{\frac{1}{1+R}\left(\alpha-\frac{\beta}{A\bar{x}+I\beta}I\alpha\right)}\right] \\
&= \mathbb{E}^i_{prior}\left[\left(\frac{A\bar{x}+I\beta}{I\beta}e^{-\frac{A\bar{x}}{A\bar{x}+I\beta}}\right)^{-\frac{\alpha}{1+R}}\right] = \mathbb{E}^i_{prior}\left[\left((1+\xi)e^{-\frac{\xi}{1+\xi}}\right)^{-\frac{\alpha}{1+R}}\right] \\
&= \left[1+\left(\left[(1+\xi)\exp\left(-\frac{\xi}{1+\xi}\right)\right]^{\frac{1}{1+R}}-1\right)\frac{\bar{\xi}}{\xi}\right]^{-\bar{\alpha}}, \qquad (4.14)
\end{aligned}$$

where

$$\xi = \frac{\bar{x}}{I}\frac{A}{\beta} = \frac{\bar{x}}{I}\frac{A}{\bar{\beta}+\sum_{k=1}^{I}N^k} > 0 \quad \text{and} \quad \bar{\xi} \equiv \frac{\bar{x}}{I}\frac{A}{\bar{\beta}}. \qquad (4.15)$$

See section B.3 in appendix B (fact B.6, p. 263) for a proof of the last step in (4.14).

The term ξ was defined in (4.7) as the *expected relative excess return*. Note that

$$\xi \equiv \frac{\mathbb{E}^i[\theta]-RP}{RP} = \frac{\alpha}{\beta}\frac{1}{RP} - 1 = \frac{\alpha}{\beta}\frac{\frac{A\bar{x}}{I}+\beta}{\alpha} - 1 = \frac{\bar{x}}{I}\frac{A}{\beta}.$$

So, ξ can be viewed as the equilibrium level of the expected excess return of the risky asset less opportunity cost, relative to those opportunity cost. Since investors can freely choose the level of information, ξ can vary between zero and $\bar{\xi}$. When the number of signals $\sum_k N^k$ in the market goes to infinity, ξ goes to zero. When, on the other hand, nobody buys signals and $\sum_k N^k = 0$, then $\xi = \bar{\xi}$ by (4.15) and the maximally feasible relative excess return is realized. So, every investor i views $\xi \in \left(0, \bar{\xi}\right]$ as inversely related to her decision variable N^i.

Intuitively, the expected relative excess return is the lower the more *common* information becomes. The reason is that more common information brings individual and average market expectations closer to each other. Since price plays a double role as an information aggregator and as opportunity cost, investors dislike this because asset holdings become less valuable.[6] However, investors also *privately* benefit from more information. As argued before, given any portfolio (b^i, x^i) the prior expected variance of the portfolio value $Rb^i + \theta x^i$ falls in N^i from the point of view of

[6]This feature of information in financial markets does not seem to be specific to fully revealing price and competitive REE. In a market maker model, Foster and Viswanathan (1996) show that profits fall the more homogeneous information becomes across investors. In another model with market makers, Jackson and Peck (1999) find examples where 'good news' make the asset price overshoot as bidding drives price *above* the level of the return. The finding also carries over to a competitive REE with partly informative price in chapter 5.

investor i, since the *ex ante* expected variance of the asset return $\mathbb{E}^i_{prior}\left[\mathbb{V}^i\left(\theta|\{S^i_1,...,S^i_{N^k}\}^I_{k=1}\right)\right]$ falls in $\sum_k N^k$. Risk-averse investors like that. It makes tomorrow's consumption less volatile. They anticipate being able to make a more educated portfolio choice once they have received signal realizations.

Using (4.14) in (4.13), *ex ante* utility becomes

$$\begin{aligned}\mathbb{E}^i_{prior}\left[U^i\right] &= -\tfrac{1+R}{R}(\delta R)^{\frac{1}{1+R}}\exp\left(-A\frac{R}{1+R}(W^i_0 - cN^i)\right) \\ &\quad\cdot\left[1+\left(\left[(1+\xi)\exp\left(-\frac{\xi}{1+\xi}\right)\right]^{\frac{1}{1+R}}-1\right)\frac{\bar{\xi}}{\xi}\right]^{-\bar{\alpha}}.\end{aligned} \tag{4.16}$$

The cost of signals cN^i enters (4.16) in the form of an initial wealth reduction. The last factor in (4.16) (on the second line) captures the effect of the relative excess return ξ on utility. The forces at work in this factor are laid out below. The term $(1+\xi)\exp(-\xi/(1+\xi))$ strictly exceeds unity iff $\xi > 0$, which is always satisfied by assumptions 4.1, 4.3 and 4.6. Hence, the last factor in (4.16) is well defined for all information levels.

Although the number of signals has to be discrete, let's pretend for now that one can take the derivative of *ex ante* utility with respect to N^i to describe the optimum. (Monotonicity of the first order condition in the relevant range will prove this to be admissible.) Differentiating (4.16) with respect to the number of signals yields the incentive to purchase information. As long as $\partial\mathbb{E}^i_{prior}\left[U^{i,*}\right]/\partial N^i > 0$, investor i will generically purchases more signals. If $\partial\mathbb{E}^i_{prior}\left[U^{i,*}\right]/\partial N^i \leq 0$ for all N^i, she purchases no information at all. Taking the derivative of (4.16) with respect to N^i, and dividing by $-\mathbb{E}^i_{prior}\left[U^i\right] > 0$ for clarity, yields

$$\begin{aligned}-\frac{1}{\mathbb{E}^i_{prior}\left[U^{i,*}\right]}\frac{\partial\mathbb{E}^i_{prior}\left[U^{i,*}\right]}{\partial N^i} &= -A\tfrac{R}{1+R}\,c \\ &\quad+\frac{\bar{\alpha}}{\bar{\beta}}\frac{\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}}\left(1-\frac{1}{1+R}\frac{\xi^2}{(1+\xi)^2}\right)-1}{1+\left(\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}}-1\right)\frac{\bar{\xi}}{\xi}}.\end{aligned} \tag{4.17}$$

The first term on the right hand side of (4.17) is negative and expresses the marginal disutility from purchasing an additional signal. It is the *marginal cost* of a signal. The second term

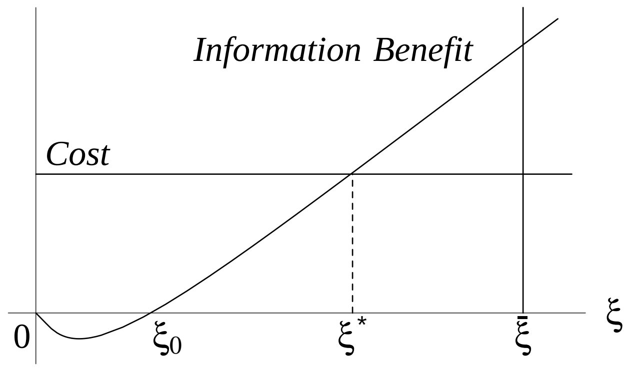


Figure 4.2: Information acquisition in equilibrium

expresses possible benefits. Let's name it the *potential marginal benefit* of a signal. It is only a potential benefit as it can turn negative when ξ drops too low. The denominator of the potential benefit term is always positive, whereas the numerator can take a negative sign for low levels of ξ. Under a strictly positive interest factor R, marginal benefit hits zero at one (and only one) point $\xi_0 > 0$, irrespective of $\bar{\xi}$ (and $\bar{\alpha}, \bar{\beta}$). These and several further properties of the marginal benefit term are presented in section B.4 of appendix B (p. 264).

Would investors ever want to acquire information under fully revealing prices? The answer is affirmative. Figure 4.2 depicts a case.[7] The potential marginal benefit curve has a long arm in the positive range that slopes strictly upward. So, as long as $\bar{\xi}$ is large enough, there is a strictly positive expected relative excess return ξ^* at which marginal benefits of a signal equal marginal cost. Although the relative excess return could attain any real value in principle, signals are not perfectly divisible. As a consequence, the precise optimal number of signals will yield a relative excess return somewhere in the neighborhood of ξ^*.

Since the expected relative excess return ξ cannot exceed $\bar{\xi}$, such an interior equilibrium can only occur if $\bar{\xi}$ is sufficiently large. Hence, investors will acquire a strictly positive amount of

[7] Parameters underlying the benefit curves in figures 4.2 through 4.5 are $A = 2$, $\bar{\alpha} = 1.3$, $\bar{\beta} = 1$, and $R = 1.1$. The level of $\bar{\xi}$ depends on average asset supply, which is $\bar{x}/I = 7$ in figures 4.2 and 4.5, $\bar{x}/I = 3$ in figure 4.3, and $\bar{x}/I = 1$ in figure 4.4. Marginal cost is given by $c = .1$, A, and R in figures 4.2, 4.3 and 4.5.

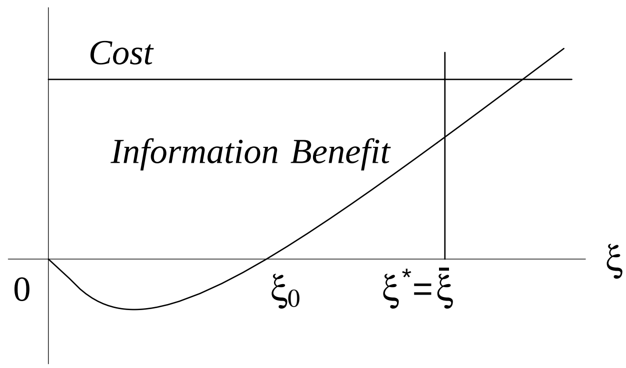


Figure 4.3: No information in equilibrium due to high signal cost

information only if the financial market meets the following two conditions. First, supply of the risky assets needs to be strong so that $\bar{x}/I$ is high. Then investors anticipate that they will invest a relatively large portion of their savings in the risky asset, and information about the risky asset return becomes relatively important to them. Second, investors need to be sufficiently risk averse relative to the mean-variance ratio of the risky asset so that $A/\bar{\beta}$ is high. Since the benefit of information stems from lowering the prior variance of the portfolio, information matters more for investors who are more risk averse.

So, the market environment determines whether information is valuable to investors indeed. Information is not a good in itself. When $\bar{\xi}$ drops too low, the marginal benefits of a signal cannot reach the point where they would meet or exceed marginal cost, and nobody will acquire a signal so that $\xi^* = \bar{\xi}$. This case is depicted in figure 4.3 (risky asset supply is reduced by more than half as compared to figure 4.2). $\bar{\xi}$ is low if relatively few risky assets are supplied to the market (low $\bar{x}/I$), or if investors are little risk averse (low A), or when the mean-variance ratio of the asset return is relatively high (high $\bar{\beta}$) so that risk matters little compared to payoff. Then investors do not value information enough to acquire it.

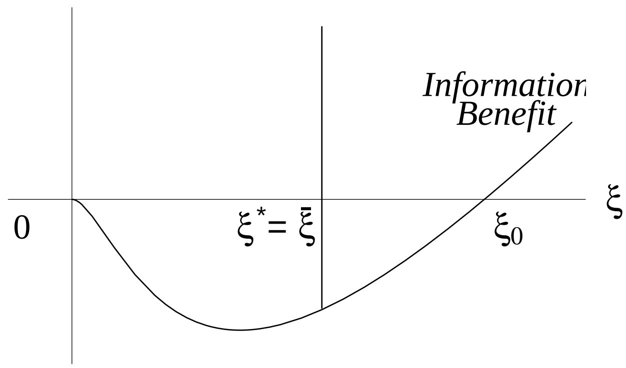


Figure 4.4: No information due to market environment

What if signal cost drops to zero? Even then, there are market conditions in which information has no or negative value. Figure 4.4 depicts a case in which the price of a signal c is zero but information would not be acquired (risky asset supply is reduced to a seventh of the level in figure 4.2). The potential benefits of a signal turn negative for low (non-zero) levels of ξ. As investors acquire more information, ξ moves away from $\bar{\xi}$ and to the west. If investors receive too much information, then what could be benefits of an additional signal turn into losses from information, and the potential marginal benefit curve dives into the negative range. Every additional signal will lower an investor's *ex ante* utility once the available amount of information has driven ξ below ξ_0. Intuitively, this happens because all private information turns public under fully revealing price and is mutually extracted to update beliefs. When the amount of acquired information is large, the negative effect from a reduced expected excess return weighs more heavily than any positive effects of more information on higher moments of the return distribution. Investors find information undesirable when it becomes too common.

The potential benefits vanish as ξ goes to zero. In this limit, no investor wants to purchase a signal. But every investor would accept signals for free. The limiting level of $\xi = 0$ is reached, for instance, when no risky assets are supplied to the market ($\bar{x}/I \to 0$). Similarly, when investors become risk neutral ($A \to 0$), or when the prior variance tends to zero ($\bar{\beta} \to \infty$), then there is no benefit of holding information but also no harm done. Finally, if investors were given infinitely many signals for free, ξ would reach zero but the return realization θ would become known with certainty and the previously risky asset would turn into a perfect substitute to the bond. The common cause for information to lose its value in all these cases is that the relative excess return is driven down to zero so that no investor chooses to hold any risky asset. In this limit, information does not have a negative value either. Investors are simply unaffected. If investors don't think at 9am that they will be holding a risky asset at 10am, they know they will never need to act upon information. An infinite amount of information makes investors indifferent to it.

Proposition 4.2 *Suppose assumptions 1 through 4.11 hold. Then a RICE (definition 4.1) has the following properties.*

- *The financial market REE is unique and symmetric so that all investors hold the same amount of risky assets $\bar{x}/I$ in equilibrium.*
- *Signals are perfect strategic substitutes.*
- *If the cost of a signal is strictly positive, then the market equilibrium for signals is unique up to a permutation of the signal allocation.*

  If the cost of a signal is nil but $R > 0$, then there are two signal market equilibria, one of which involves an infinite amount of freely received signals.
- *Investors acquire a strictly positive and finite number of signals in a signal market equilibrium if and only if $\bar{\xi}$ is sufficiently large.*

Proof. Under assumptions 1 through 4.5 asset price is fully revealing (corollary 4.1.1). Since asset demand is $x^{i,*} = (\beta/A)(\mathbb{E}^i[\theta] - RP)/RP$ by (4.6), every investor holds the same amount of risky assets $x^{i,*} = \beta\xi/A = \bar{x}/I$ by (4.15). This establishes symmetry of the financial REE.

Investor i's own number of signals N^i enters the first order condition (4.17) in the same way as any other investor j's signal acquisition, viz. through $\sum_{k=1}^{I} N^k$. Thus, signals are perfect strategic substitutes.

For $c > 0$, uniqueness of $\sum_{k=1}^{I} N^{k,*}$ in equilibrium follows from the fact that the positive arm of the marginal benefit term in (4.17) is strictly monotonically increasing in ξ. See lemma B.1 in section B.4 of appendix B (p. 264). Since the marginal cost of an additional signal is constant and strictly positive, there is a unique intersection of the marginal cost and marginal benefit curve. Call the unique level of ξ at which the curves intersect ξ^*. If $c = 0$, there is a second equilibrium at $\xi = 0$, in which $\sum_{k=1}^{I} N^k \to \infty$.

If $\xi^* \geq \bar{\xi}$, the unique information equilibrium entails no information acquisition and $\xi^* = \bar{\xi}$. As $\bar{\xi}$ increases, there will be a unique information equilibrium with exactly one acquired signal since the marginal benefit term in (4.17) is strictly monotonically increasing in ξ. As $\bar{\xi}$ moves further up, there will be a new and unique information equilibrium with exactly two acquired signals for the same reason, and so forth. So, investors acquire a strictly positive amount of signals if, and only if, $\bar{\xi}$ is sufficiently large. Only the equilibrium level of ξ is uniquely determined (but is generically different from ξ^*). So, the sum $\sum_k N^k$ is unique but the equilibrium assignment of signals to investors can be altered freely.

With these arguments, the marginal analysis that lead to (4.17) is retroactively proven to be valid because (4.17) is strictly monotonically increasing in ξ over the relevant range $\xi \in (\xi_0, \infty)$.

■

With a fully revealing price, signals are perfect strategic substitutes. That is, my fellow investors' signal is as good or bad as my own one because its realization gets revealed through equilibrium price together with all other signal realizations. As a consequence, the equilibrium does not determine how many signals a single investor holds. In equilibrium, one investor may acquire all $\sum_i N^i$ signals while nobody else buys any signal, or all investors may hold the same number of signals. The equilibrium number of signals results in a relative excess return in the neighborhood of ξ^*. The value of information about the risky asset is closely linked to the value of the competing riskless bond.

Corollary 4.2.1 *Under the conditions of proposition 4.2, the following is true for a RICE.*

- *For any $R \in (0, \infty)$, there is a $\hat{\xi}$ so that at least one signal is acquired in equilibrium if $\bar{\xi} \geq \hat{\xi}$ and no information is acquired if $\bar{\xi} < \hat{\xi}$.*
- *In the limit when $R \to \infty$, an information market equilibrium involves no information acquisition if signals are costly ($c > 0$).*
- *For $R = 0$, the benefits of information are strictly positive at any ξ. Then, if $c = 0$, there is a unique information market equilibrium which involves infinite information acquisition.*

Proof. The first statement is an immediate corollary to the proof of proposition 4.2. The second statement follows since the numerator in the marginal benefit term in (4.17) vanishes for $R \to \infty$. For $R = 0$, the marginal information benefit cannot drop below zero by claim B.2 in section B.4 of appendix B (p. 265). So, there is only one equilibrium if $c = 0$, proving the third statement. ∎

There is always a market size, or a degree of risk aversion, or a level of the mean-variance ratio of the risky asset so that information becomes worthwhile to acquire in equilibrium. The only exception is the degenerate limiting case where the interest rate of the bond becomes infinite (and the potential benefit curve coincides with the horizontal axis). When the bond becomes entirely worthless ($r = -1, R = 0$), investors do not want to hold it in their portfolio. In this extreme case, they would choose to acquire an infinite amount of information about the risky asset return as signal costs fall to zero. Intuitively, if there is no riskless asset in the economy, investors want to create a riskless asset by acquiring infinitely much information about a risky asset.

So, information can be worthwhile to acquire in the present framework even under fully revealing price. However, information need not be desirable. From a different perspective, corollary 4.2.1 clarifies again that information can turn from a public good into a public bad as market conditions change. These market conditions are captured in $\bar{\xi}$ and can be affected by R. In financial markets, information is a tertiary commodity. Investors are concerned about consumption, the primary good. Assets are mere means to the end of consumption, or secondary commodities. Information, finally, has value only if it helps investors make better portfolio decisions with regard to these assets. In this sense, information is a tertiary commodity. As such it can change its character under different conditions.

4.4 Informational Efficiency and Informativeness of Price

In the present framework, alternative efficiency concepts for information in financial markets can be compared. One can treat information just as any other economic commodity and apply a Pareto criterion.

Definition 4.2 (Informational Pareto efficiency) *An allocation of $x^{i,**}$ risky assets, $b^{i,**}$ riskless bonds, and $N^{i,**}$ signals to investors $i = 1, ..., I$ is called informationally Pareto efficient in a given market environment if there is no other allocation such that all investors are at least as well off and at least one investor is strictly better off.*

It does not matter for this Pareto criterion that information can change from a public good into a public bad. The criterion is conditional on a given market environment. To investigate whether the RICE in section 4.3 is Pareto efficient, think of a benevolent social planner who can dictate every consumer j to buy exactly $N^{j,**}$ signals. This social planner maximizes $\sum_{j=1}^{I} \mathbb{E}^{i}_{prior}\left[U^{j}\right]$ with respect to $\{N^1, ..., N^I\}$. Thus, similar to Samuelson's (1954) condition for public good provision, a benevolent social planner's first order conditions for information allocation are not (4.17) but, rather,

$$
\begin{aligned}
&-\frac{1}{\mathbb{E}^{j}_{prior}\left[U^{j,**}\right]}\frac{\partial \sum_{k=1}^{I}\mathbb{E}^{k,**}_{prior}\left[U^{k,**}\right]}{\partial N^{k}} = -A\tfrac{R}{1+R}c \\
&+\frac{\bar{\alpha}}{\bar{\beta}}\frac{\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}}\left(1-\frac{1}{1+R}\frac{\xi^2}{(1+\xi)^2}\right)-1}{1+\left(\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}}-1\right)\frac{1}{\xi}\frac{\bar{x}A}{I\bar{\beta}}}\left(1+\sum_{k\neq j}^{I}\frac{\mathbb{E}^{k}_{prior}\left[U^{k,**}\right]}{\mathbb{E}^{j}_{prior}\left[U^{j,**}\right]}\right)
\end{aligned}
\tag{4.18}
$$

for any $j \in 1,...,I$, written in terms of that investor j's utility. Thus, compared to the privately perceived benefits, the potential benefits that a social planner considers in his decision are scaled up by a factor of $1 + (1/\mathbb{E}^{j}_{prior}\left[U^{j,**}\right]) \cdot \sum_{k\neq j}^{I}\mathbb{E}^{k}_{prior}\left[U^{k,**}\right] > 1$. Therefore, if information is a public bad, a benevolent social planner wants to implement an even smaller amount of information than the private market. However, since no information is acquired in private markets in that case anyway, the market equilibrium is informationally efficient when information is a public bad.

On the other hand, if information is a public good under given market conditions, a social planner wants (weakly) more information to be allocated than markets provide. Individual investors do not take into account that their signal acquisition also benefits other investors through fully revealing price. In this case, markets allocate (weakly) less information than desirable. However, since signals are not divisible, one cannot infer from condition (4.18) that a social planner wants to implement strictly more information. It could happen theoretically that an additional signal pushes relative excess return ξ down so much that all investors are worse off and not better. So, only a weak efficiency statement can be made, which holds up to discrete tolerance. In figure 4.5, a social planner wants to allocate information so that relative excess return is brought down from around ξ^* to ξ^{**}. However, if an additional signal makes the implementable level of ξ drop far below ξ^{**}, investors are better off if relative excess return remains at the market equilibrium level around ξ^*.

Proposition 4.3 *Suppose assumptions 1 through 4.11 hold. Then the following is true in a RICE (definition 4.1).*

- *If $\bar{\xi} \leq \xi_0$, then the equilibrium is informationally Pareto efficient.*

- *If $c > 0$ and at least one signal is acquired in equilibrium, then the equilibrium is not informationally Pareto efficient up to discrete tolerance.*

- *If $c = 0$, then, unless $R = 0$, the equilibrium with finite information is informationally Pareto efficient, whereas the equilibrium with infinitely much information is not Pareto efficient.*

Proof. First, if $\bar{\xi} \leq \xi_0$, information benefits are weakly negative by lemma B.1 (appendix B, p. 264) and a social planner would not allocate any signal. Second, if $c > 0$ and at least one signal is acquired in equilibrium, then the equilibrium level of ξ (around ξ^*) must be strictly lower than $\bar{\xi}$, and the marginal benefit term in (4.17) must be strictly positive. Then the augmented marginal benefit term of the social planner in (4.18) must strictly exceed marginal cost at the equilibrium level of ξ^*. Up to discrete tolerance, increasing the number of signals by one augments the sum of investors' *ex ante* utilities.

Third, if $c = 0$, then the marginal benefit term in (4.17) must be as close to zero in equilibrium as possible because investors must have chosen a number of signals such that ξ is as close to zero or ξ_0 as possible. So, the two equilibria are *locally efficient* in the sense that a small change in ξ would violate the Pareto criterion. However, the equilibrium with finite information always yields higher utility than the equilibrium with infinite information. Since $c = 0$, signal choice only affects the last factor in *ex ante* utility (4.16), that is the term (4.14). By claim B.1 (appendix B, p. 264), this term is strictly decreasing in ξ for $\xi \in (0, \xi_0)$ (it equals $h(\xi)^{-\bar{\alpha}}$). Since CARA utility is negative, utility must be strictly increasing until ξ reaches ξ_0. So, only the equilibrium with a finite number of signals can be informationally Pareto efficient. ■

When information is for free and $c = 0$, only the market outcome with finite information is efficient but not the one with infinite information. In other words, as long as the bond is valuable ($R > 0$), neither markets nor the social planner want perfect common information. Investors prefer having a second asset around that is not a perfect substitute to the bond. That means that the second asset has to be risky. Risk-averse investors don't love risk but they do like to hold risky

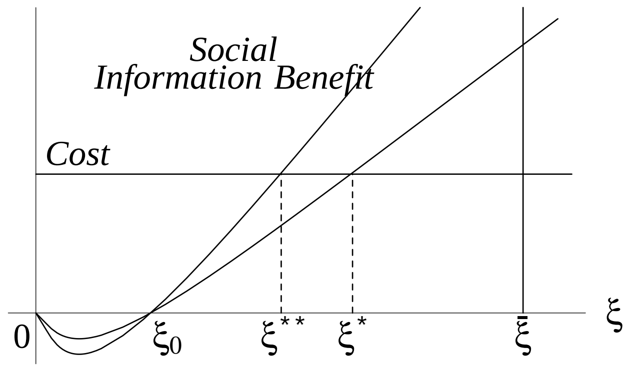


Figure 4.5: Socially desirable information choice

assets, as long as those assets yield a positive excess return over opportunity cost. Only once the bond lost all value and $R = 0$ ($r = -1$), investors prefer to remove all risk from the single remaining asset by receiving infinite information.[8]

The informational efficiency of financial markets can be, and has been, judged with further criteria. Fama discerns three degrees of market efficiency: (i) A financial market is called strong-form efficient if no investor or group of investors has monopolistic access to information. (ii) Semi-strong efficiency is satisfied if prices incorporate publicly available information. (iii) Weak efficiency demands only that investors' information sets contain historical prices. Since prices are fully revealing, the RICE in the present framework satisfies strong-form efficiency. To make Fama's (1970) criteria operational, the variance of the price is often used as a measure inversely related to its informativeness (see e.g. Holden and Subrahmanyam (1996), Foster and Viswanathan (1996), Back, Cao and Willard (2000); however, Easley and O'Hara (1992) apply a measure of entropy instead). It is a special property of the normal distribution that the posterior variance is deterministic. This is not the case for Poisson signals.

[8]Burguet and Vives (2000) find that unbounded information acquisition can be prevented only if the marginal cost of information is strictly positive. In the present model, unbounded information acquisition can be prevented even if $c = 0$. Moreover, bounded information Pareto dominates unbounded information strictly.

In general, a statistically well defined measure of the informativeness of a signal is its precision, the reciprocal of its prior expected variance. It becomes

$$\frac{1}{\mathbb{E}^i_{prior}\left[\mathbb{V}^i\left(P|\theta\right)\right]} = \frac{\left(\frac{A\bar{x}}{I}+\bar{\beta}+\sum_{i=1}^I N^{i,*}\right)^2}{\mathbb{E}^i_{prior}\left[\sum_{i=1}^I \cdot N^{i,*}\theta\right]} = \frac{\bar{\beta}}{\bar{\alpha}}\frac{\left(\frac{A\bar{x}}{I}+\bar{\beta}+\sum_{i=1}^I N^{i,*}\right)^2}{\sum_{i=1}^I N^{i,*}}$$

for equilibrium price, a Poisson variable, by (4.6) and fact B.5 (appendix B, p. 262). So, the precision of price can fall with the number of signals purchased! Since

$$\frac{\partial}{\partial N^i}\left(\frac{1}{\mathbb{E}^i_{prior}\left[\mathbb{V}^i\left(P|\theta\right)\right]}\right) = -\frac{\bar{\beta}}{\bar{\alpha}}\frac{\left(\frac{A\bar{x}}{I}+\bar{\beta}+\sum_k N^{k,*}\right)\left(\frac{A\bar{x}}{I}+\bar{\beta}-\sum_k N^{k,*}\right)}{\left(\sum_{i=1}^I N^{i,*}\right)^2},$$

each additional signal reduces the precision of the market clearing price if the amount of pre-existing information $\sum_{k=1}^I N^{k,*}$ is small. An additional signal improves precision if and only if $\sum_{k=1}^I N^{k,*}$ is larger than $A\bar{x}/I+\bar{\beta}$.

This may seem paradoxical at first but it really is not. Each investor anticipates that she and all others will respond to signals in their portfolio choice. Investors will no longer base their decision on priors only. So, from an *ex ante* perspective, asset demand (4.6) becomes more volatile with more information, given any market clearing price. The anticipated variance of asset demand is

$$\mathbb{E}^i_{prior}\left[\mathbb{V}^i\left(x^{i,*}|\theta\right)|RP\right] = \frac{\bar{\alpha}}{\bar{\beta}A^2(RP)^2}\left(\sum_{k=1}^I N^k\right)$$

by fact B.5. However, financial markets need to clear. So, every investor ends up holding $\bar{x}/I$ risky assets in equilibrium by proposition 4.2, irrespective of what her information is. Hence, market price has to absorb fully any demand moves that stem from information revelation. As a consequence, the variance of price can increase with more information acquisition. When there is relatively little pre-existing information $\sum_{k=1}^I N^{k,*}$, an additional signal will affect individual demands strongly and thus add to the price's variance. If, on the other hand, a lot of information is available already, an additional signal that gets fully revealed through price will move investors' demands little. If investors receive many signals, an additional piece of information is likely to confirm previous observations and tends to stabilize demand. So, equilibrium price is expected to become less volatile with an additional signal if the pre-existing information level $\sum_{k=1}^{I} N^{k,*}$ is high.[9]

Proposition 4.4 *In a RICE under assumptions 1 through 4.11, the* ex ante *precision of the price system decreases with every additional signal if and only if* $\sum_{k=1}^{I} N^{k,*} < A\bar{x}/I + \bar{\beta}$.

Rational investors completely internalize this change in price volatility when they maximize their *ex ante* utility. In that sense, the precision of price is not directly related to welfare. However, a Pareto criterion based on investors' utilities does reflect the social value of information.

4.5 Conclusion

How much information do investors buy, and how much should they buy? To address this question, a rational expectations equilibrium is considered in which both asset markets at Wall Street and information markets for private detective services clear. This equilibrium is compared to a social planner's preferred allocation. To make the framework concrete, a gamma distribution of the asset return is employed, along with Poisson distributed signals and CARA utility.

On a first stage, investors choose how many signals to buy, given the choice of signals of all other investors. The individually optimal number of signals maximizes each investor's *ex ante* (expected indirect) utility, based on prior knowledge, and clears the market for signals. On a second stage, investors learn the realizations of the signals they purchased, and decide on their portfolio and consumption given these signal realizations. Asset supply and demand are equalized through market price. For simplicity, investors are assumed to be price takers on both stages. Prices are fully revealing under an appropriate set of assumptions.

Grossman and Stiglitz' (1980) paradox that no equilibrium exists if the asset price is fully revealing can be resolved in the present framework. More information reduces the *ex ante* variance

[9]The response of the variance of price to private information seems to be model specific. Grossman and Stiglitz (1980) conjecture that "the more individuals who are informed, the more informative is the price system" but cannot confirm this because positive and negative effects are mutually offseting in their model, and informativeness of price remains constant. Verrecchia (1982) confirms the conjecture in the competitive REE of his model under a partially revealing price.

of the asset return from the point of view of the individual investor. So, it allows investors to make a more educated portfolio choice and reduces the expected variance of future consumption. Consequently, more information increases *ex ante* utility of risk averse investors. This gives every investor an individual incentive to acquire information or to remain uninformed if information is too costly. A comprehensive equilibrium in both the financial market and the market for information results.

In a rational expectations equilibrium, investors buy a positive amount of information whenever markets are large enough, when investors are sufficiently risk averse, or when the variance of the risky asset is relatively high compared to its payoff. Due to the public good nature of information, this equilibrium is informationally inefficient from a benevolent social planner's point of view. However, irrespective of the price of a signal, information can turn into a public bad if markets are small, investors not very risk averse, or the variance of the return low. Then, additional information only has a negative effect. The reason is that a sufficient statistic, summarizing all investors' private information, becomes known to everybody through fully revealing price. So, beliefs become more homogeneous across agents, and consequently market price moves closer to every individual investor's return expectations. This diminishes the value of the risky asset from the point of view of each individual investor and can outweigh positive effects of information.

The following chapter 5 carries out further analyses within the present framework. Shortcomings notwithstanding, the normal distribution function can be a convenient choice both for the return distribution and the signal distribution. Most importantly, the normal-normal conjugate prior distribution permits a theoretical investigation into *partly* revealing asset prices. Under simplifying assumptions, a closed-form equilibrium at least at Wall Street results. Concluding remarks on this line of research follow at the end of chapter 5.

Chapter 5

Towards a theory of information acquisition in financial markets

The previous chapter left questions open: What consequences do less than fully revealing but still informative asset prices have? What are the incentives for information acquisition under different assumptions on the asset and signal distributions? The present chapter explores directions for generalizations.

It remains true under partly informative prices that a risk averse investor likes information. Information sharpens her knowledge, and this allows her to make a more adequate portfolio choice. Anticipating this improved choice, her *ex ante* utility rises because a less risky portfolio choice means safer consumption tomorrow. However, information also continues to have a bad side in financial markets. If information is widely disseminated, the expected excess return of an asset over its opportunity cost falls. In general, prices play a double role: They reflect the opportunity cost of an asset, and they aggregate and disseminate information to everybody. It is this double role of common information that can harm investors in financial markets. If more information gets to the market, this information is at least partly transmitted through price. But then, when rational investors update their information, their expectation of the dividend gets closer to market expectations. In other words, the excess return of an asset over its opportunity cost falls with more information.

In the case of gamma distributed asset returns, this negative effect prevailed if markets were small or returns not very risky, and no information was acquired in these circumstances and under the assumptions of chapter 4. However, the negative effect of common information was outweighed in large markets and for relatively risky assets. With a normally distributed asset return and under fully revealing prices, this negative effect is so strong that no investor ever wants to buy any information. However, choosing the normal distribution for the asset return can be an attractive assumption when modelling the more complicated case of a partly informative price, in which some external noise remains. Then, investors do have incentives to acquire information.

5.1 Generalizations

For the most general model, without any assumptions on the distribution of the asset return, I prove the following in the present chapter: If signals are costly, then an investor acquires no information if she is risk neutral. A risk neutral investor is indifferent between a riskless bond and a risky asset and would not act on information when it arrives. So, information has no value to a risk neutral investor and she would not pay for it.

The general model is hard to understand when neither distributions nor the utility function are specified. To make the model tractable for extensions, I assume again that utility is CARA but make all random variables Gaussian. When the asset price is fully revealing, Grossman and Stiglitz' conjecture is again proven to fail in a particularly surprising way: No investor wants to obtain information, not even receive it for free, when prices at Wall Street are fully revealing. The distributional assumptions of the model in this chapter are identical to those of Grossman and Stiglitz. As a consequence, a unique equilibrium both at Wall Street and at the news stands exists in which no information is acquired. The profound reason is that information may lose the character of a good and turn into a public bad when it becomes too common. This will also be the recurring insight when prices are only partly, and not fully, informative. Information can have features of a negative externality in financial markets. Fully revealing prices are merely an extreme case.

What value does information have when prices are only partially revealing? Hellwig (Hellwig 1980) posed the same question early on and built a general model similar to the present one. That model, however, does not have a closed-form solution. I aim to obtain a closed-form solution for my extension and make two additional assumptions. First, each investor has to choose the membership in either of two groups. She can either become a news watcher and do what the group representative mandates, or become a price watcher. This will affect the equilibrium definition. Second, I make the additional assumption that all signals are sold in perfect copies. Information comes only from newspapers, so to say. This is not so unrealistic considering that the large majority of investors obtains information from publicly accessible media in practice. The assumption rules out, however, that an investor flies to the stock-issuing country or talks in private to the president of the stock-issuing firm.

It can be shown in this model that information still reduces the excess return even if prices are only partially revealing. From the point of view of less informed investors (price watchers), information behaves like a negative externality. Well informed investors (news watchers) feel the positive impact of a lowered variance, but still suffer from the loss in the excess return. This loss is so strong that a symmetric equilibrium, in which all investors are equally well informed, cannot exist. The model provides a much needed closed-form solution for the financial market equilibrium at Wall Street (but not at the newsstands), and becomes tractable. However, there is no simple way to assess the incentives of a news watcher or price watcher to open a third (one-person) group with different information. As a consequence, the existence of an according equilibrium may be compromised. This issue can be resolved in simulations of the defection case in future research.

The model in this chapter corroborates the reasoning of chapter 4 why individuals have an incentive to obtain information. More information can raise *ex ante* utility for a risk averse individual. In general, *ex ante* utility is increasing when the expected excess return of the risky asset increases. Moreover, it tends to increase when the variance of the portfolio falls. So, information is good because it sharpens our knowledge about the dividend and thus tends to reduce the variance of the portfolio. However, information also has a bad side in financial markets since it affects the excess return negatively. The excess return of an asset can be viewed as its expected dividend less the opportunity cost of acquiring it. In the notation to be adopted again soon, the excess return may be defined as $\mathbb{E}[\theta] - RP$, where θ is the dividend, R the yield of a riskless bond, and P the price of the risky asset.

In general, prices play a double role: They reflect the opportunity cost of an asset, and they aggregate and transmit information. This double role is precisely why more information can harm investors in financial markets. If more information gets to the market, this information is, at least partly, transmitted through prices. But then, when rational investors update their information, their expectation of the dividend gets closer to market expectations which are incorporated in the price. In other words, the excess return $\mathbb{E}[\theta] - RP$ is likely to be falling with more information! The equilibrium price P and the expected relative payoff $\mathbb{E}[\theta/R]$ get closer to each other. Under fully revealing prices, this effect is so strong that no investor wants to buy any information, and even a benevolent social planner agrees that no information is the right choice. Under partly informative prices, in which some external noise remains, this effect is still present. From the point of view of less informed investors, information behaves like a negative externality. Well informed investors feel the positive impact of a lowered variance, but still suffer from the loss in the excess return. This loss is so strong that a symmetric equilibrium, in which all investors are equally well informed, cannot exist.

5.2 The Lucas Tree Model with Information Acquisition

As in chapter 4, suppose again that all securities lose their value after one period. In addition, suppose that there are but two assets on Wall Street: One riskless bond and one risky stock. When Wall Street opens today at 10am, investors can choose their portfolio. Both assets will yield a payoff tomorrow once and for all. The bond is going to pay the principal plus interest $R = 1 + r$ tomorrow, whereas the stock is going to pay a risky dividend θ. Today, the bond costs exactly one dollar, while the stock will go for P dollars to be set by a Walrasian auctioneer at 10am. All investors hold prior beliefs about the distribution of the dividend θ. Newsstands open at 9am today.

This is just like the Lucas tree model of a financial market, reduced to two periods. It is almost identical to the commonly used model of terminal wealth maximization, just that investors also decide about consumption in the present period. The main innovation of the models in this dissertation is the addition of a second market, a market for information.

Investors can, of course, choose to completely ignore newspapers and only observe security prices at Wall Street. I call them *price watchers*. Price watchers know that the price P will convey market information about tomorrow's dividend since P is a function of all other investors' asset demands which in turn reflect their information. If at least some of the other investors have purchased newspapers, the pure price watcher can free-ride on their information by merely looking at the price. In the most extreme case, the stock price at 10am will *fully reveal* all investors' information (not θ, of course, but the content of all newspapers that others have read). Chapter 4 only considered this possibility. Section 5.3 revisits it. Alternatively, the price may contain noise. Then it only *partly informs* about other investors' knowledge—a more realistic possibility analyzed at large in section 5.4. A pure price watcher combines his prior knowledge about θ with the information that he can extract from the price, and then makes his portfolio choice.

However, investors may also choose to buy newspapers at 9am. I call investors who do so

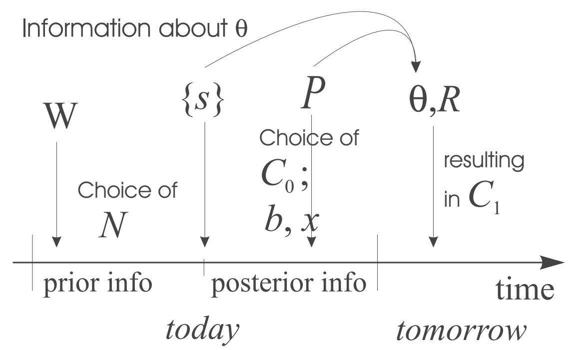


Figure 5.1: Timing of information revelation and decisions

news watchers. Besides reading newspapers, news watchers still use the price P to extract additional information (unless they consider it redundant to the information in the newspapers). To become a news watcher requires a fixed but not sunk cost F for wasting time with the sales person at the news stand, for reading the newspaper, and for taking time to interpret the information.[1] Each newspaper sells at a unit cost c.

How many different newspapers should a news watcher buy? When standing in front of the news stand at 8.55am, each news watcher knows that she will base her portfolio decision, to be taken at 10am today, on the information that she is about to get out of the newspapers at 9am. She also knows the statistical distribution of the information in the newspaper, which is more informative than her own prior beliefs. Taking all this into account, she rationally evaluates what a newspaper is worth to her and makes her best choice. Formally, a news watcher maximizes her *ex ante* indirect utility with respect to the number of newspapers—given her information at 8.55am, her anticipation of how signals are likely to affect her beliefs in five minutes, and her expected portfolio choice to be taken at 10am. Figure 5.1 replicates figure 4.1) in the previous chapter to illustrate the timing of decisions again.

[1]The fact that F is fixed and not sunk allows each news watcher to become a price watcher by buying $N^i = 0$ newspapers. I will also consider the case of $F = 0$.

Given her wealth W_0^i and her prior information, a news watcher first chooses the number of signals (newspapers) N^i to buy. A news watcher then receives the realizations $\{s_1^i, ..., s_{N^i}^i\}$ of these N^i signals $\{S_1^i, ..., S_{N^i}^i\}$ (she gets to know the newspaper content). Given this information, she finally chooses consumption today, C_0^i, and decides how many bonds b^i and how many risky x^i assets to hold for consumption tomorrow.

To make things concrete, let investor i maximize additively separable utility under a discount factor $\delta < 1$ and an instantaneous utility function $u(\cdot)$ that is increasing and concave. That is, let her maximize

$$U^i = \mathbb{E}\left[u(C_0^i) + \delta u(C_1^i) \left| \mathcal{F}^i \right.\right] \tag{5.1}$$

with respect to consumption today, C_0^i, and tomorrow, C_1^i, and a portfolio choice. $\mathcal{F}^i$ denotes the information set available to investor i at the time of her portfolio choice. Besides the price P, it contains the realizations $\{s_1^i, ..., s_{N^i}^i\}$ of the N^i signals $\{S_1^i, ..., S_{N^i}^i\}$ that she acquired. For ease of notation, I will usually abbreviate investor i's conditional expectations $\mathbb{E}\left[\cdot \left| \mathcal{F}^i \right.\right]$ with $\mathbb{E}^i[\cdot] \equiv \mathbb{E}\left[\cdot \left| \mathcal{F}^i \right.\right]$. These expectations are different for each investor in general unless *all* information is publicly available and commonly known. I do not make any assumptions on the asset distribution at this point.

The intertemporal budget constraint of investor i is

$$b^i + Px^i = W_0^i - C_0^i - F^i - cN^i \tag{5.2}$$

so that

$$C_1^i = Rb^i + \theta x^i \tag{5.3}$$

will be available for consumption tomorrow. The investor is endowed with initial wealth W_0^i, and decides about her consumption C_0^i and C_1^i in each period, her holdings of the riskless bond b^i, her holdings of the risky stock x^i, and how much information N^i she wants to buy. If an investor acquires at least one newspaper, she has to incur the fixed cost F. To indicate this, I use the shorthand $F^i \equiv \mathbf{1}(N^i \geq 1) \cdot F$. While assets are assumed to be perfectly divisible, signals have to be acquired in discrete amounts.[2]

On the second stage, after having received the realizations of her N^i signals $\{s^i_j\}_{j=1}^{N^i}$, the investor decides on asset holdings and consumption given the asset price P. A price watcher receives no signals and simply relies on the price. In any case, at the second stage every investor has updated his or her beliefs about the dividend's distribution to a posterior distribution, given the signal realizations. The Euler conditions for the problem at this stage are therefore

$$\frac{1}{\delta} = R\,\mathbb{E}^i\left[\frac{u'(C_1^{i,*})}{u'(C_0^{i,*})}\right] \tag{5.4}$$

and

$$\frac{P}{\delta} = \mathbb{E}^i\left[\theta\,\frac{u'(C_1^{i,*})}{u'(C_0^{i,*})}\right], \tag{5.5}$$

where expectations $\mathbb{E}^i[\cdot]$ are conditional on the realizations of the signals and the asset price. The optimal choices $C_1^{i,*}$, $b^{i,*}$ and $x^{i,*}$ are decision rules depending on the price P, on the chosen number of signals N^i (which was decided earlier), and on the information transmitted through the signal realizations and the price P. The choices of $C_1^{i,*}$, $b^{i,*}$ and $x^{i,*}$ imply a level of posterior indirect utility, which we can denote by $U^{i,*} = u(C_0^{i,*}) + \delta\mathbb{E}^i\left[u(C_1^{i,*})\right]$.

On the first stage, the investor chooses the number of signals she wants to receive. She does this by maximizing *ex ante* utility given her information before the realizations of the signals arrive. At this time she cannot know more than the prior parameters of the respective distributions, but she builds her new *ex ante* beliefs by taking into account how signals will most likely change her beliefs in the next stage. *Ex ante* utility is $\mathbb{E}^i_{pre}\left[U^{i,*}\right] = \mathbb{E}^i_{pre}\left[u(C_0^{i,*})\right] + \delta\mathbb{E}^i_{pre}\left[u(C_1^{i,*})\right]$ by the law of iterated expectations. The optimal number of signals $N^{i,*}$ maximizes *ex ante* utility $\mathbb{E}^i_{pre}\left[U^{i,*}\right]$.

These observations immediately imply

Lemma 5.1 *Suppose signals are costly. Then an investor acquires no information*

[2] For a signal to contain information, its distribution has to depend on θ. So, a continuum of signals (or an infinite number of them), will a.s. reveal the exact realization of θ to news watchers. For markets to clear, P must equal θ/R in this case, otherwise news watchers want to reshuffle their portfolio. But then the price fully reveals θ itself and removes all uncertainty—an unrealistic case of little interest.

- *if she is risk neutral, or*
- *if the prior distribution is insensitive to changes in the number of signals. That is, if the prior distribution of the asset return coincides with the* ex ante *distribution that reflects the rationally anticipated use of a signal.*

Proof. Suppose the investor is risk neutral. Then there is no benefit from a signal. *Ex ante* utility degenerates to $C_0^{i,*} + \delta \mathbb{E}_{pre}^i \left[C_1^{i,*} \right]$. For a risk neutral investor to neither demand a positively nor a negatively infinite number of assets, $\mathbb{E}^i [\theta] = RP$ and $R = 1/\delta$ in a financial market with no arbitrage (or in equilibrium). Thus, *ex ante* utility becomes $C_0^{i,*} + \delta \mathbb{E}_{pre}^i \left[C_1^{i,*} \right] = W_0^i - F^i - cN^i$ by (5.2) and (5.3). *Ex ante* utility of a risk neutral investor is independent of the portfolio composition. As a result, signals only cause costs, but do not have a benefit, which proves the first statement. To prove the second statement, suppose the prior distribution of the fundamental is insensitive to the number of signals. Then an additional signal weakly reduces both $u(C_0^{i,*})$ and $u(C_1^{i,*})$, and strictly reduces at least one of the two, for *any* future realization of the dividend. Since the prior distribution of the fundamental is supposed not to change, a signal cannot have a benefit in this case either. ∎

A risk neutral investor is indifferent whether she holds a risky stock or a riskless bond in her portfolio. Hence, she would never act upon information, which makes information useless to her. As immediate as lemma 5.1 may seem, it has important consequences. It clarifies that the incentive to purchase costly information is closely linked to risk aversion and higher-order moments of the risky asset's distribution. Risk averse investors do care about their portfolio composition, whereas their risk neutral colleagues don't. The fact that information has no value for risk neutral investors also highlights that information is not a good or bad in its own right. It has only value if it affects our decisions.

For the remainder of this chapter, I make the following assumptions, which are identical to those of chapter 4.

Assumption 5.1 (Common risk aversion) *Investors are risk averse and share a common and certain degree of risk aversion, all else equal.*

Assumption 5.2 (Common priors) *Investors hold the same prior beliefs about the distributions of the risky asset return, the signals, and the supply of the risky asset.*

Assumption 5.3 (Finite Risk) *The prior variance of the risky asset return is strictly positive and finite.*

Assumption 5.4 (No borrowing constraint) *Investors can carry out unlimited short sales.*

However, the following assumption differs from assumption 4.5 in chapter 4. It allows for the possibility that the supply of the risky asset may be uncertain. While section 5.3 will consider a fixed supply as in chapter 4, section 5.4 will drop that assumption and make x uncertain.

Assumption 5.5 (Market size) *The number of investors I is discrete and known. The (expected) supply of the risky asset $\bar{x}$ is strictly positive.*

The following assumptions are identical again to those of chapter 4.

Assumption 5.6 (Exogenous asset and signal supply) *Supply of the risky asset $\bar{x}$ is finite and strictly positive. Supply of the riskless asset and supply of the signals are perfectly elastic.*

Assumption 5.7 (Unique information) *All signals $\{S_1^i, ..., S_{N^i}^i\}_{i=1}^I$ are conditionally independent given the realization of the asset return, $S_n^i|\theta \overset{i.i.d.}{\sim} f(s_n^i\,|\theta)$.*

Assumption 5.8 (Equal precision of signals) *All signals have equal precision given the realization of the asset return.*

Assumption 5.9 (Price taking) *Investors are price takers in all markets.*

Assumptions 5.1 through 5.5 are made for convenience. Together with a fixed asset supply they also happen to be necessary conditions for prices to become fully revealing. However, an uncertain asset supply will ultimately prevent prices from being fully revealing (section 5.4). As Hellwig (1980) observed, assumption 5.9 stands in a certain conflict to investor rationality. Investors are assumed not to take into account how their asset demand affects price. Yet, they are assumed to perceive how the equilibrium price correlates with their own information through their demand. Hellwig called investors of this kind schizophrenic. Kyle (1989) provided a way out by allowing that investors behave like monopsony firms when demanding assets. However, to enhance tractability of the model, I retain assumption 5.9 throughout this chapter. This is the subject for future generalizations. Finally, to obtain closed-form solutions both for fully *and* for partly revealing asset prices, let's suppose the following.

Assumption 5.10 (CARA) *Investors have CARA utility with* $u(C) = -e^{-AC}$.

Assumption 5.11 (Normality) *Random variables are Gaussian.*

This last assumption implies that the dividend realization can be negative or positive. Consequently, it entails the more profound assertion that investors are prevented from rejecting a negative payoff through a well-working legal system. This is an unfortunate weakness of the Gaussian model that the gamma-Poisson model as in chapter 4 did not exhibit. Since the normal distribution is a conjugate prior to itself, assumption 5.11 implies conditional independence (assumption 5.7). When all signals have identical variance σ_S^2, assumption 5.8 is satisfied.

5.3 Fully Revealing Prices

In this section, I reconsider the benchmark case of fully revealing prices both to revisit Grossman and Stiglitz's (1980, 1980) famous paradox and no-equilibrium conjecture, and to clarify the bolts of the model. For this, I make the assumption that supply of the risky asset is certain and known to all investors. It takes the value $\bar{x}$. I will first establish the financial market equilibrium at Wall Street. Second, I will prove the existence of a unique rational expectations equilibrium both at Wall Street and at the news stands under fully revealing prices. I finally discuss its efficiency properties.

5.3.1 The financial market equilibrium

Suppose that, at 10am, all investors have possibly different information about the two parameters μ^i and τ^i of the risky asset's distribution. Dividends are normally distributed. So, $\theta \sim \mathcal{N}\left(\mu^i, (\tau^i)^2\right)$. At 8.55am, however, all investors have been assumed to share the same priors about the distribution of θ (assumption 5.2). So, $\mu^i_{prior} = \bar{\mu}_\theta$ and $\tau^i_{prior} = \bar{\tau}_\theta$. We would like the distribution of the signals to be such that both the prior and the posterior distribution of θ are normal. Formally, we want the dividend's distribution to be a conjugate prior distribution to the signals' distribution. In fact, assuming a normal distribution of the signals (assumption 5.11) does the job. Concretely, let each signal be independently normally distributed conditional on θ with $S^i_j|\theta \sim \mathcal{N}(\theta, \sigma^2_S)$. Then we obtain

Fact 5.1 *Suppose that the prior distribution of θ is a normal distribution with given mean $\bar{\mu}_\theta$ and variance $\bar{\tau}^2_\theta$. Suppose also that the signals $S^i_1, ..., S^i_{N^i}$ are independently drawn from a normal distribution with unknown mean θ and conditional variance σ^2_S . Then the posterior distribution of θ, given the realizations $s^i_1, ..., s^i_{N^i}$ of the signals, is a normal distribution with a mean-variance ratio*

$$\frac{\mu^i}{(\tau^i)^2} = \frac{\bar{\mu}_\theta}{\bar{\tau}^2_\theta} + \frac{1}{\sigma^2_S}\sum_{j=1}^{N^i} s^i_j$$

and variance

$$(\tau^i)^2 = \frac{1}{\frac{1}{\bar{\tau}^2_\theta} + \frac{1}{\sigma^2_S}N^i}.$$

Proof. Apply fact C.1 in appendix C.1 (p. 267) to conditionally independent signals. ∎

The mean-variance ratio $\mu^i/(\tau^i)^2$ will play an important role for investors' decisions. The posterior mean μ^i can be inferred by multiplying the mean-variance ratio with $(\tau^i)^2$. Since a sum of normal variables is normally distributed, fact 5.1 implies that the *ex ante* expectation of the posterior mean is $\mathbb{E}^i_{pre}\left[\mu^i\right] = \bar{\mu}_\theta$. It is independent of the number of signals as it has to be in general by the law of iterated expectations. While the posterior mean is a random variable, the normal-normal pair of distributions has the rare property that the posterior variance $(\tau^i)^2$ is certain given the chosen number of signals. In light of lemma 5.1, it is important that the *ex ante* variance is changing in the number of signals N^i. Indeed, fact 5.1 is good news for risk averse individuals: News watchers can lower the *ex ante* variance of the risky asset $\mathbb{E}^i_{pre}\left[(\tau^i)^2\right] = (\tau^i)^2 = 1/\left(\frac{1}{\tau_\theta^2} + \frac{1}{\sigma_S^2}N^i\right)$ by purchasing more information.

For now, let's only focus on the financial market equilibrium, restrict attention to the equilibrium at Wall Street and disregard the market for newspapers for a moment. So, we have taken a time jump to 10am. For CARA utility, the marginal utility ratios in first order conditions (5.4) and (5.5) become $u'(C_1^i)/u'(C_0^i) = e^{-A(C_1^i - C_0^i)}$. Since

$$C_1^i - C_0^i = (1+R)b^i + (\theta + P)x^i - W_0^i + F^i + cN^i$$

by (5.2) and (5.3), the first order conditions (5.4) and (5.5) simplify to

$$\frac{1}{\delta} = R\,\mathbb{E}^i\left[e^{-A(C_1^i - C_0^i)}\right] = R \cdot H^i\,\mathbb{E}^i\left[e^{-Ax^i\cdot\theta}\right] \quad (5.6)$$

$$\frac{P}{\delta} = \mathbb{E}^i\left[\theta\, e^{-A(C_1^i - C_0^i)}\right] = H^i\,\mathbb{E}^i\left[\theta \cdot e^{-Ax^i\cdot\theta}\right] \quad (5.7)$$

where $H^i \equiv \exp\left(-A\left[(1+R)b^i + Px^i - W_0^i + F^i + cN^i\right]\right)$ is certain. The expected values in (5.6) and (5.7) have simple closed-form solutions for a normally distributed dividend. They are reported as facts C.2 and C.3 in appendix C.1 (p. 267). Applying these facts to (5.6) and (5.7), and dividing one by the other, yields demand for the risky asset

$$x^{i,*} = \frac{1}{A}\,\frac{\mathbb{E}^i\left[\theta\right] - RP}{(\tau^i)^2} \quad (5.8)$$

with $\mathbb{E}^i\left[\theta\right] = \mu^i$. As is well known, demand for the risky asset is independent of wealth for CARA utility. Throughout this chapter, the term $\mathbb{E}^i\left[\theta - RP\right]$ in (5.8) will be key. It denotes investor i's expected excess return of the risky asset over the opportunity cost of one unit of the risky asset. A news watcher will go short in the risky asset whenever $\mathbb{E}^i\left[\theta\right] = \mu^i < RP$, that is whenever her posterior expectation of the dividend falls short of opportunity costs RP, and go long otherwise.

In equilibrium, asset supply equals asset demand, that is $\sum_{i=1}^{I} x^{i,*} = \bar{x}$, where $\bar{x}$ is assumed to be certain for the purposes of the present section. Thus, the equilibrium price P of the risky asset is implicitly given by

$$RP = \frac{1}{\frac{1}{I}\sum_{i=1}^{I}\frac{1}{(\tau^i)^2}}\left[\left(\frac{1}{I}\sum_{i=1}^{I}\frac{\mu^i}{(\tau^i)^2}\right) - \frac{A\bar{x}}{I}\right]. \tag{5.9}$$

This relationship sheds light on the double role of prices in financial markets. On the one side, RP is the opportunity cost of one unit of the risky asset, indicating its scarcity or value to investors. On the other side, prices aggregate all investors' information. Neglecting $\bar{x}$, RP can also be viewed as market expectations of the risky asset return (where market expectations are the average expected dividend, weighted by subjective variances). Looking back at (5.8), we could also have stated *cum grano salis* that a news watcher will go short in the risky asset whenever her posterior expectation of the dividend falls short of market expectations RP, and go long otherwise. This already hints at the fact to be established later that investors will reduce their asset demand in situations in which their own information is very similar to the market information.

Given optimal asset demand $x^{i,*}$, *posterior* indirect utility of investor i can be shown to equal

$$U^{i,*} = -\tfrac{1+R}{R}(\delta R)^{\frac{1}{1+R}}\ e^{-A\frac{R}{1+R}(W_0^i - F^i - cN^i)}\ \mathbb{E}^i\left[e^{-Ax^{i,*}(\theta - RP)}\right]^{\frac{1}{1+R}} \tag{5.10}$$

for CARA utility, irrespective of the distribution of the risky asset return (see appendix C.2, p. 269). Only investors who buy a positive amount of signals have to pay the fixed cost F. To express this, I have used the short hand $F^i \equiv \mathbf{1}(N \geq 1)\cdot F$ in (5.10) again. For a normal distribution of the dividend, the last factor in (5.10) becomes

$$\mathbb{E}^i\left[\exp\left\{-Ax^{i,*}(\theta - RP)\right\}\right] = \exp\left\{-\frac{1}{2}\left(\frac{\mu^i - RP}{\tau^i}\right)^2\right\}$$

by fact C.2 and asset demand (5.8). Note that RP is certain from a posterior point of view. Using this in (5.10), posterior indirect utility for a normally distributed dividend can be written

$$U^{i,*} = -k^i \cdot \exp\left\{A\frac{R}{1+R}(F^i + cN^i)\right\} \cdot \exp\left\{-\frac{1}{2}\frac{(\tau^i)^2}{1+R}\left(\frac{\mu^i - RP}{(\tau^i)^2}\right)^2\right\} \tag{5.11}$$

for $k^i \equiv \frac{1+R}{R}(\delta R)^{\frac{1}{1+R}} \exp\left\{-A\frac{R}{1+R}W_0^i\right\} > 0$.

The key term in (5.11) is

$$\tau^i \, \frac{\mu^i - RP}{(\tau^i)^2} = \frac{\mu^i - RP}{\tau^i}.$$

In the preceding chapter, I referred to a very similar term as *expected relative excess return* (ξ on p. 113). However, the variance τ^i appears in the above term so we would have to call it an *expected excess-return-standard-deviation ratio*. But for the remainder of this dissertation, I will refer to $\tau^i\,(\mu^i - RP)/(\tau^i)^2$ simply as the *key term*. The key term reflects both desires of a risk averse investors. For one, she wants a possibly high excess return over opportunity cost. The larger the difference $\mu^i - RP$, the better for her. On the other hand, she also wants a possibly low variance of her portfolio since she is risk averse. The lower τ^i the better. This also hints at one of the most important insights to be established soon: investors may suffer a utility loss when their own beliefs get very similar to the market information so that excess return drops. If the positive effect on variance is small, information obtained by news watchers and transmitted to others through price may in fact have features of a negative, and not a positive externality.

Investors $i = 1, ..., I$ choose their portfolios, b^i and x^i, given their respective information sets $\mathcal{F}^i = \{RP; s_1^i, ..., s_{N^i}^i\}$ if they are news watchers and $\mathcal{F}^i = \{RP\}$ if they are price watchers. If prices are fully revealing, however, then the average of all signal realization received by any investor will become known to everyone through the price. So, if prices are fully revealing, all information sets become the same $\mathcal{F}^i = \mathcal{F}$.[3]

[3] Formally, all investors' information sets become $\mathcal{F}^i = \mathcal{F} = \{\frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{N^i} s_j^i\}$. A complete derivation of this result can be found in appendix C.9 (p. 282). There the fully revealing equilibrium is treated as a special case of the most general model.

Then fact 5.1 implies that the equilibrium price P of the risky asset is implicitly given by

$$RP = \frac{1}{\frac{1}{\bar{\tau}_\theta^2} + \frac{1}{\sigma_S^2}\frac{1}{I}\sum_{i=1}^{I} N^i}\left(\frac{\bar{\mu}_\theta}{\bar{\tau}_\theta^2} + \frac{1}{\sigma_S^2}\frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{N^i} s_j^i - \frac{A\bar{x}}{I}\right), \tag{5.12}$$

where RP is the opportunity cost of the risky asset. In equilibrium, every investor chooses an optimal number of signals, given the information choice of all other investors. So, $\sum_{k \neq i} N^k$ is known to every investor i. Since everything else in RP but the sum of signal realizations is also known to every investor, the sufficient statistic $\frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{N^i} s_j^i$ becomes fully revealed to everyone through price. The following lemma explicitly restates necessary conditions for this to occur.

Lemma 5.2 *If the asset return and the signals are Gaussian, the equilibrium price of the risky asset P fully reveals all market participants' information $\frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{N^i} s_j^i$ if only if*

- *assumptions 5.1 through 5.5 are satisfied,*
- *supply of the risky asset is certain,*
- *the total number of all other investors' signals $\sum_{j \neq i}^{I} N^i$ is known to each investor i at the time of the portfolio choice, and*
- *assumption 5.8 is satisfied.*

Proof. Necessity of assumptions 5.1 through 5.3, 5.5 and 5.8 follows by inspection of the general solution for the market price

$$RP = \frac{1}{\frac{1}{I}\sum_{i=1}^{I}\frac{1}{A^i}\left(\frac{1}{\bar{\tau}_\theta^2} + \frac{1}{\sigma_S^2}N^i\right)}\left(\frac{1}{I}\sum_{i=1}^{I}\frac{1}{A^i}\left(\frac{\bar{\mu}_\theta}{\bar{\tau}_\theta^2} + \frac{1}{\sigma_S^2}\sum_{j=1}^{N^i} s_j^i\right) - \frac{\bar{x}}{I}\right).$$

Similarly, inspection shows that the risky asset supply needs to be certain. For the necessity of assumption 5.4 and sufficiency, compare the arguments in corollary 4.1.1 (p. 115). ∎

Lemma 5.2 restates all conditions that appeared in corollary 4.1.1 for price to be fully revealing under a gamma-Poisson conjugate pair. In addition to those conditions, lemma 5.2 also requires that signals be equally precise, which is not implied by the choice of the normal-normal conjugate pair.

In the light of lemma 5.2, fully revealing asset prices seem unlikely to occur in practice. They are still an important theoretical benchmark case. Yet, ever since Grossman and Stiglitz's (1980) seminal article, fully revealing market prices have been dismissed with the theoretical argument that no equilibrium existed. In this and related arguments, some important features of informational externalities appear to have been overlooked as results in the following subsection may clarify.

5.3.2 The information market equilibrium

A complete market equilibrium both at Wall Street and the news stands can be defined as follows.

Definition 5.1 Rational Information Choice Equilibrium) *A rational information choice equilibrium (RICE) is an allocation of $x^{i,*}$ risky assets, $b^{i,*}$ riskless bonds, and $N^{i,*}$ signals to investors $i = 1, ..., I$. It involves an asset price P, a signal price c and a fixed cost of news watching F along with a set of consistent beliefs such that*

1. *the portfolio $(x^{i,*}, b^{i,*})$ s optimal given opportunity cost RP and the respective posterior information sets $\mathcal{F}^i$ for investors $i = 1, ..., I$*
2. *the market for the risky asset clears, $\sum_{i=1}^{I} x^{i,*} = \bar{x}$, and*
3. *the choice of signals $N^{i,*}$ is optimal for investors $i = 1, ..., I$ given the sum of all other investors' signal choices $\sum_{j \neq i} N^{j,*}$, and given the costs c and F.*

Just as with the RICE definition 4.1 in chapter 4, the mixture of elements of a Bayesian Nash equilibrium with those of a Walrasian equilibrium makes the definition 'hybrid.' On the first stage, investors choose the number of signals given the choice of all other investors, and a Bayesian Nash equilibrium results. Since information is a kind of public commodity both under partly and under fully revealing priced, each investor needs to take everybody else's choice into account. On the second stage, investors are price takers. That is, they consider the impact of their demand on equilibrium price as negligible, and a competitive Walrasian REE results given the Bayesian Nash equilibrium on the first stage. One could call this a hybrid equilibrium since investors behave as price takers on the second change but rationally anticipate on the first stage that each investor's small contribution to price makes it fully revealing in the aggregate (investors are also price takers with respect to signal cost).[4]

The market for the riskless bond clears by the assumption of a perfectly elastic world supply given the world interest factor $R = 1 + r$. The market for information clears under the implicit assumption that there is an infinitely elastic supply of information. That is, any number of signals can be produced at unit cost c. Given their anticipation of a financial market equilibrium as outlined in the previous subsection, investors choose their level of information on the first stage. Since the cost of becoming a news watcher is fixed, but not sunk, a choice of $N^i = 0$ signals means that an investor decides to become a price watcher.

To make her choice of information, each investor maximizes *ex ante* indirect or *ex ante* utility. If prices are fully revealing, the posterior parameters are the same for all investors, that is $\mu^i = \mu$ and $\tau^i = \tau$ for all investors i. Therefore, $RP = \mu - \frac{A\bar{x}}{I}\tau^2$ for fully revealing prices by (5.9) and posterior utility is certain—a striking difference to the gamma-Poisson model in the preceding chapter. Hence, *ex ante* utility $\mathbb{E}^i_{pre}\left[U^{i,*}\right]$ simply equals the certain posterior indirect utility

$$\mathbb{E}^i_{pre}\left[U^{i,*}\right] = -k^i \exp\left\{A\frac{R}{1+R}(F^i + cN^i)\right\} \cdot \exp\left\{-\frac{1}{2}\frac{\tau^2}{1+R}\left(\frac{A\bar{x}}{I}\right)^2\right\} \tag{5.13}$$

for

$$\tau^2 = \frac{1}{\frac{1}{\bar{\tau}^2_\theta} + \frac{1}{\sigma^2_S}\frac{1}{I}\sum_{k=1}^{I} N^k}.$$

Even though investors may only choose a discrete number of signals, it does no harm in the present context if we differentiate *ex ante* utility (5.13) with respect to N^i. Taking the derivative and multiplying it by the positive factor $-(1+R)/\mathbb{E}^i_{pre}\left[U^{i,*}\right]$ yields

$$-\frac{1+R}{\mathbb{E}^i_{pre}\left[U^{i,*}\right]}\frac{\partial \mathbb{E}^i_{pre}\left[U^{i,*}\right]}{\partial N^i} = -AR\,c - \frac{1}{2}\left(\frac{A\bar{x}}{I}\right)^2\frac{\tau^4}{I\sigma^2_S} < 0. \tag{5.14}$$

[4]See the discussion of the 'schizophrenia' critique in chapter 4, p. 116.

So, no matter whether investors add a discrete or real number of signals, each additional signal lowers their *ex ante* utility! Therefore, the unique rational expectations equilibrium involves zero information. No investor wants to acquire any signal even if nobody else acquires a signal. Moreover, even if newspapers were free of charge ($c = F = 0$), investors would refuse to open them and throw them away unread.

Proposition 5.1 *Suppose that assumptions 5.1 through 5.11 hold and that asset price is fully revealing. Then there is a unique rational expectations equilibrium. No investor acquires information in this equilibrium even if signals are for free.*

Proof. By inspection of (5.13). ■

The profound reason for this result is that the key term is reduced by more information for all investors. As argued before, the key term reflects both desires: the one for a high excess return and the one for a low variance. For a normally distributed dividend, however, the desires can never be met through more publicly available information. The key term becomes

$$\frac{\mu^i - RP}{\tau^i} = \tau \left(\frac{\mu}{\tau^2} - \frac{\mu - A\bar{x}\tau^2/I}{\tau^2} \right) = \frac{A\,\bar{x}\,\cdot\,\tau}{I}$$

for a fully revealing price. In other words, since τ shows up in the numerator only, we must conclude that the negative impact of more information, as it reduces excess return, always dominates the positive effect of a lowered variance. More information is always bad for investors when dividends are normally distributed and asset price is fully revealing.

Since investors in Grossman and Stiglitz' (1980) model only have a choice between one signal or no signal, their model is a special case of the present framework which allows for any finite number of signals to be acquired ($N^i \in \mathbb{N}_0$). As quoted in the preceding chapter, too, Grossman and Stiglitz (1980, *Conjecture 6*) wrote: "In the limit, when there is no noise, prices convey all information, and there is no incentive to purchase information. Hence, the only possible equilibrium is one with no information. But if everyone is uninformed, it clearly pays some individual to become informed. Thus, there does not exist a competitive equilibrium." This and similar conjectures can be found in the literature ever since. Recent examples include Romer (1993) and Barlevy and Veronesi (2000). The latter authors remark: "Finally, as Grossman and Stiglitz point out, we need to prevent prices from being fully revealing; otherwise an equilibrium will fail to exist." This no-equilibrium conjecture is proven to be wrong in the present more general framework. The reason is that, even if everyone is uninformed, it does not pay any individual to become informed. Under fully revealing prices, any signal reduces the expected excess return of the risky asset $\mathbb{E}^i\left[\theta - RP\right] = \frac{A\bar{x}}{I}\tau$. Consequently, information is not a public good, but a public bad under fully revealing prices.

Somewhat separately, a series of articles investigated fully revealing equilibria, too (see e.g. Allen 1981, Jordan 1982, Rahi 1995, Pietra and Siconolfi 1998). These articles establish that, generically, a fully revealing rational expectations equilibrium exists at Wall Street. Beyond those articles, the present framework explicitly analyzes the incentives for information acquisition and incorporates a market for information at the news stands. The present framework presents an example in which a fully revealing equilibrium exists but prices will contain no information. One might conjecture that the result is less extreme under alternative distributional assumptions. This is the case in fact. Chapter 4 showed for the more realistic assumption of a gamma distributed dividend that investors may acquire information even under fully revealing prices. If investors are highly risk averse, they start buying information about a gamma distributed asset return even under fully revealing prices.

The key difference between the normal-normal and the gamma-Poisson conjugate pair is that posterior utility is certain under the normal-normal pair whereas it is uncertain under the gamma-Poisson pair. The reason is that the variance of the asset return is certain in the normal-normal case. In the gamma-Poisson case, on the other hand, an unfavorable signal (low $\bar{\theta}$) has a high precision and a favorable signal (high $\bar{\theta}$) has a low precision. As a consequence, uncertainty about the precise effect on demand and price remains and information can have a value even under fully revealing asset prices. In the gamma-Poisson case, investors like to buy more signals to improve their information when markets satisfy certain conditions. To the contrary in the normal-normal case, acting on a signal acquisition moves price strongly and closer to expected return. For the normal-normal conjugate pair, the negative effect of more information always outweighs any individual benefit.

5.3.3 Informational efficiency

As in chapter 4, a complete financial and information market equilibrium does exist under fully revealing prices. However, now investors choose to acquire no information at all so that nothing can get revealed in fact. Is this informationally efficient?

Like in chapter 4, the present consumption maximization framework allows for a classical welfare analysis, applied to information. Think of a benevolent social planner who can dictate every investor i to buy exactly $N^{i,**}$ signals. This social planner maximizes $\sum_{i=1}^{I} \mathbb{E}_{pre}^{i}\left[U^{i,**}\right]$ with respect to $\{N^1, ..., N^I\}$, where $U^{i,**}$ denotes posterior indirect utility after the social planner has implemented an allocation of newspapers to investors.[5] Thus, similar to Samuelson's (1954) seminal condition for public good provision, a benevolent social planner's does not consider condition (5.14) for signal acquisition but rather

$$\begin{aligned} -\frac{1+R}{\mathbb{E}_{pre}^{i}\left[U^{i,**}\right]} \frac{\partial \sum_{k=1}^{I} \mathbb{E}_{pre}^{k,**}\left[U^{k}\right]}{\partial N^{i}} &= -ARc \\ &\quad -\frac{1}{2}\left(\frac{A\bar{x}}{I}\right)^{2} \frac{\tau^{4}}{I\sigma_{S}^{2}} \left(1+\sum_{k\neq i}^{I} \frac{\mathbb{E}_{pre}^{k}\left[U^{k,**}\right]}{\mathbb{E}_{pre}^{i}\left[U^{i,**}\right]}\right) \end{aligned} \tag{5.15}$$

for $k = 1, ..., I$ (pretending again that signals are divisible for simplicity's sake). Thus, the second term in every investor's condition is scaled up by a factor of $1+(1/\mathbb{E}_{pre}^{k}\left[U^{k,*}\right]) \cdot \sum_{i\neq k}^{I} \mathbb{E}_{pre}^{i}\left[U^{i,*}\right] > 1$ for every single investor i. Since information is not beneficial but undesirable for each and every individual investor under fully revealing prices, the benevolent social planner emphatically agrees with the private market solution: No news watchers under fully revealing prices, please. If one price watcher bought a signal and became a news watcher, he would not only reduce his own excess return

[5] Passing by, I have confined the social planner to leaving every investor with his or her beliefs. The social planner cannot transfer knowledge between investors.

$\mathbb{E}^i[\theta - RP]$, but that of any other investor, too. The unique rational expectations equilibrium is therefore informationally efficient.

Proposition 5.2 *Suppose that assumptions 5.1 through 5.11 hold and that asset price is fully revealing. Then the unique rational expectations equilibrium, in which no investor acquires information, is informationally efficient.*

Proof. By inspection of the sum of (5.13), or (5.15). ∎

Proposition 5.2 also sheds some new light on Grossman and Stiglitz' (1980) more general assertion that financial markets are unlikely to be informationally efficient in general. Information need not have the character of a public good in all circumstances. It may actually be a public bad, and no information acquisition can be socially desirable. Under fully revealing prices, information is a perfect strategic substitute. No matter which price watcher dares to buy a signal, he harms himself and everybody else in the market.

5.4 Partly Informative Prices

The previous section showed that investors rationally choose not to become news watchers when prices are fully revealing. To arrive at a more realistic information market equilibrium, suppose that the asset price is only partly informative about tomorrow's dividend. For this, only one of the necessary conditions in lemma 5.2 needs to fail. To make things concrete and to keep the tradition of the previous literature, suppose that supply of the risky asset is uncertain. To keep things simple, assume that investors cannot buy information about asset supply. In particular, let the asset supply be normally distributed with $X \sim \mathcal{N}(\bar{x}, \omega_x^2)$ and independent of any other random variable in the model.

The derivation of the financial market and information market equilibrium is based on an extension of Hellwig (1980). Whereas Hellwig's general model does not have a closed-form solution, I aim to obtain a closed form for my extension and make two additional assumptions. First, each investor has to choose the membership in either of two groups. She can either become a news watcher and do what the group representative mandates, or become a price watcher. This assumption will affect the equilibrium definition and is therefore not labelled for now. No closed-form solution in the financial market exists for more than two groups of investors. Second, I make an additional assumption.

Assumption 5.12 (Perfect copies) *All signals are sold in perfect copies.*

This is not so unrealistic considering that the large majority of investors obtains information from publicly accessible media in practice. Assumption 5.12 rules out, however, that an investor may personally sniff around at the headquarters of a company, send out a private detective, or talk in private to the chief-executive officer at the stock-issuing firm.

I proceed in similar steps as in the previous section. First, I derive the complete financial market equilibrium in closed form and discuss, second, its immediate implications for information acquisition. Third, I analyze necessary properties of the information equilibrium, which does not have a closed-form solution. The equilibrium value of information needs to be such that no group member has an incentive to deviate. The exact no-deviation conditions remain to be established in future research.[6]

5.4.1 The financial market equilibrium

Investors can update their information both through newspapers and through observation of the asset price. Then investors act upon these signal realizations and signal realizations make their way into asset price. Therefore, one signal, the asset price, is no longer conditionally independent of the other signals. As a consequence, to derive a complete financial and information market equilibrium under partly informative prices, we need a generalization of fact 5.1 to the case of

[6] The precise conditions have to be derived in the absence of closed-form solutions even in the financial market since a deviating investor can form a (third) group of his own.

correlated signals. The according property is reported as fact C.1 in appendix C.1 (p. 267). Rational investors, who know the correlation in equilibrium, update their beliefs accordingly. They infer a conditional distribution of θ—given the signal realizations that they received, given the equilibrium price that they observed, and given their respective correlation as it occurs in equilibrium.

After correctly inferring the correlation between their signals and opportunity cost RP in equilibrium, rational investors base their portfolio choice on this knowledge. Thus, a rational expectations equilibrium as in definition 5.1 is a fixed point that results in no excess asset demands and consistent beliefs (see Hellwig 1980 for a general argument). Due to the mutual dependence of asset demands on equilibrium beliefs and beliefs on equilibrium asset demands, a complete financial market and information market equilibrium can be complicated to characterize and often has no closed-form solution.[7]

To obtain a closed-form solution for the financial market equilibrium in this section, I consider a subclass of equilibria. As before, there are two groups of investors: Price watchers and news watchers. Now, however, I require that news watchers be a homogeneous group. They must not independently decide on different amounts of information. Instead, they must jointly pick a number of identical newspapers, acquire them and read them or not. A *news watcher representative* enters an agreement with all news stands at 8.55am to offer exactly N different newspapers and to sell one copy of each to every news watcher at 9am. If the group representative determines that a strictly positive number of newspapers be purchased, all news watchers agree to go to a news stand at 9am, to pay the fixed cost F and to buy N newspapers at a cost of c each. If the group representative happens to mandate that no newspaper be purchased, news watchers jointly become price watchers and do not pay the fixed cost F. Among the I investors, a share $\lambda \equiv I^{NW}/I$ chooses to be news watching in equilibrium.

An according equilibrium definition is

[7] For the derivation of the equilibrium under fully revealing prices in the previous section, we were able to take a shortcut at this point. In particular, we could use the fact that the information sets of all investors had to coincide under fully revealing prices. As a result, we never needed to explicitly consider the correlation of signals and RP (but investors implicitly evaluated this correlation correctly). A detailed derivation can be found in appendix C.9 (p. 282).

Definition 5.2 (Two-Group Rational Information Choice Equilibrium) *A two-group RICE is an allocation of $x^{i,*}$ risky assets and $b^{i,*}$ riskless bonds to investors $i = 1, ..., I$, a share λ of news watchers, and an allocation of N signals to each news watcher. It involves an asset price P, a signal price c and a fixed cost of news watching F along with a set of consistent beliefs such that*

1. *the portfolio $(x^{i,*}, b^{i,*})$ is optimal for every investor $i = 1, ..., I$ given opportunity cost RP and the respective information set $\mathcal{F}^{NW}$ for a news watcher and $\mathcal{F}^{PW}$ for a price watcher*
2. *the market for the risky asset clears, $\sum_{i=1}^{I} x^{i,*} = \bar{x}$,*
3. *(a) the choice of N signals is optimal for every news watcher given that there are λI news watchers, and given the costs c and F, and*

   (b) receiving no signal is optimal for every price watcher given that there are λI news watchers receiving N signals.

Condition 3 is the main requirement in this definition. First of all, in equilibrium a news watcher must not want to object to the group representative about the choice of N. Or, in other words, she must want to read the N newspapers that she is required to buy and not want any further newspaper. Similarly, a price watcher must not have an incentive to switch group. If $N^* = 0$ or $\lambda^* = 0$ or both, everybody is a price watcher in equilibrium.

In the previous section on fully revealing prices we have seen that the equilibrium asset price (5.12) is a linear function of the signals $\sum_i \sum_j s_j^i$ and the certain asset supply $\bar{x}$. Now, the supply X of the risky asset is uncertain and all news watchers buy copies of the same N newspapers (assumption 5.12). Yet, suppose that there is a unique financial market equilibrium under partly informative prices, in which the price will satisfy a very similar linear structure. Suppose,

$$RP = \pi_0 + \pi_S \sum_{j=1}^{N} S_j - \pi_X X \tag{5.16}$$

for three coefficients π_0, π_S, π_X to be determined. That this guess is right will be confirmed soon.

To make his portfolio choice at 10am, each price watcher takes into account how θ and RP are jointly distributed from a posterior perspective. At this time, he extracts all possible information from his observation of RP and infers the most likely realization of the dividend θ applying fact C.1 (appendix C.1, p. 267). To update his beliefs to posterior beliefs, a price watchers departs from his *ex ante* knowledge. At 9am he knows that there are λI news watchers and that they read N newspapers. So, from a price watcher's perspective, the joint *ex ante* normal distribution of θ and RP has a vector of means $\bar{\mu}^{PW} = (\bar{\mu}_\theta; \pi_0 + \pi_S N \bar{\mu}_\theta - \pi_X \bar{x})^T$ and a variance-covariance matrix

$$\bar{\Sigma}^{PW} = \begin{pmatrix} \bar{\tau}_\theta^2 & \pi_S N \bar{\tau}_\theta^2 \\ \pi_S N \bar{\tau}_\theta^2 & \pi_S^2 N \left(N \bar{\tau}_\theta^2 + \sigma_S^2 \right) + \pi_X^2 \omega_X^2 \end{pmatrix}.$$

Recall that signals are conditionally normally distributed $S_j|\theta \sim \mathcal{N}(\theta, \sigma_S^2)$. Statistically, this implies $\mathbb{V}(S_j) = \mathbb{V}_\theta(\mathbb{E}[S_j|\theta]) + \mathbb{E}_\theta[\mathbb{V}(S_j|\theta)] = \bar{\tau}_\theta^2 + \sigma_S^2$.

When Wall Street opens, the price watcher observes RP, updates his *ex ante* to posterior beliefs applying fact C.1, and arrives at the updated expected value of the dividend

$$\mathbb{E}[\theta | RP; \lambda, N] = \mu^{PW} = m_0^{PW} + m_{RP}^{PW} RP \tag{5.17}$$

and the updated variance of the dividend $\mathbb{V}(\theta | RP; \lambda, N) = (\tau^{PW})^2$, where

$$m_0^{PW} = \frac{(\pi_S^2 N \sigma_S^2 + \pi_X^2 \omega_X^2)\bar{\mu}_\theta - \pi_S N (\pi_0 - \pi_X \bar{x}) \bar{\tau}_\theta^2}{\pi_S^2 N (N \bar{\tau}_\theta^2 + \sigma_S^2) + \pi_X^2 \omega_X^2}, \tag{5.18}$$

$$m_{RP}^{PW} = \frac{\pi_S N \bar{\tau}_\theta^2}{\pi_S^2 N (N \bar{\tau}_\theta^2 + \sigma_S^2) + \pi_X^2 \omega_X^2}, \tag{5.19}$$

$$(\tau^{PW})^2 = \frac{(\pi_S^2 N \sigma_S^2 + \pi_X^2 \omega_X^2) \bar{\tau}_\theta^2}{\pi_S^2 N (N \bar{\tau}_\theta^2 + \sigma_S^2) + \pi_X^2 \omega_X^2}. \tag{5.20}$$

A news watcher proceeds in a similar manner. Given any choice of N that the news watcher group happens to take, she considers the *ex ante* joint normal distribution of θ, RP, and the N signals. Then she asks herself, what her posterior knowledge will be, once having received the signal realizations $s_1, ..., s_N$ and having observed RP. For this, she can take into account that nobody else will receive better information than she does. Other investors are either price watchers and receive no signal at all, or they are news watchers and receive exact copies of her own N signals. As a consequence, prices are fully redundant for her. Prices contain no additional information beyond the knowledge that she gets out of her N newspaper copies already. A formal proof of the redundancy of RP is given in appendix C.3 (p. 270).

Therefore, a news watcher can disregard RP for her updating and simply apply fact 5.1 (p. 145). As a result, her posterior belief about the dividend is that it is normally distributed with conditional mean

$$\mathbb{E}\left[\theta \,|RP; s_1, ..., s_N; \lambda, N\right] = \mu^{NW} = m_0^{NW} + m_S^{NW} \sum_{j=1}^{N} s_j \tag{5.21}$$

and conditional variance $\mathbb{V}\left(\theta \mid RP; s_1, ..., s_N; \lambda, N\right) = (\tau^{NW})^2$, where

$$m_0^{NW} = \frac{\sigma_S^2 \bar{\mu}_\theta}{\sigma_S^2 + \bar{\tau}_\theta^2 N}, \tag{5.22}$$

$$m_S^{NW} = \frac{\bar{\tau}_\theta^2}{\sigma_S^2 + \bar{\tau}_\theta^2 N}, \tag{5.23}$$

$$(\tau^{NW})^2 = \frac{\sigma_S^2 \bar{\tau}_\theta^2}{\sigma_S^2 + \bar{\tau}_\theta^2 N}, \tag{5.24}$$

by fact 5.1.

We now know the subjective posterior distributions of all investors. Investors base their portfolio decisions on these posterior distributions, and demand $x^{i,*}$ as given by (5.8) for $i = PW, NW$. Asset markets at Wall Street must clear. So,

$$(1-\lambda) \cdot x^{PW,*} + \lambda \cdot x^{NW,*} = \frac{x}{I},$$

where x is the realization of the uncertain asset supply X. Hence, the realization of equilibrium price must satisfy

$$
\begin{aligned}
RP &= \frac{1}{(1-\lambda)\frac{1-m_{RP}^{PW}}{(\tau^{PW})^2}+\lambda\frac{1}{(\tau^{NW})^2}} \\
&\quad \left((1-\lambda)\frac{m_0^{PW}}{(\tau^{PW})^2}+\lambda\frac{m_0^{NW}}{(\tau^{NW})^2}+\lambda\frac{m_S^{NW}}{(\tau^{NW})^2}\sum_{j=1}^{N}s_j - A\frac{x}{I}\right) \\
&= \frac{1}{\frac{1}{\bar{\tau}_\theta^2}+\left[(1-\lambda)\frac{\pi_S(\pi_S N-1)}{\pi_S^2 N\sigma_S^2+\pi_X^2\omega_X^2}+\lambda\frac{1}{\sigma_S^2}\right]N} \\
&\quad \left(\frac{\bar{\mu}_\theta}{\bar{\tau}_\theta^2}-(1-\lambda)\frac{\pi_S N(\pi_0-\pi_X\bar{x})}{\pi_S^2 N\sigma_S^2+\pi_X^2\omega_X^2}+\lambda\frac{1}{\sigma_S^2}\sum_{j=1}^{N}s_j - A\frac{x}{I}\right).
\end{aligned}
\tag{5.25}
$$

The second step follows from (5.18) through (5.20) and (5.22) through (5.24). We can now match the coefficients π_0, π_S, and π_X in equation (5.16) with the according terms in (5.25). This yields a non-linear equation system in three equations and the three unknowns π_0, π_S, π_X. The equation system happens to have a unique closed-form solution.

Lemma 5.3 *Suppose that assumptions 5.1 through 5.11 hold, that assumption 5.12 is satisfied and that asset supply is uncertain and Gaussian. Then there exists a unique two-group financial market equilibrium for a given share λ of news watchers and a given number of signals N under equilibrium definition 5.2.*

Proof. The closed-form solution of this equilibrium is derived in appendix C.4, p. 271. Uniqueness can be established by assuming price to be a higher-order functional of $\sum_{j=1}^{N} S_j$ and X, and leading that assumption to a contradiction. ■

This financial market equilibrium is still a partial equilibrium, given that there are λI news watchers who purchase N signals each. Our main focus lies on its implications for the incentives to acquire information and the simultaneous information market equilibrium at the news stands.

A first insight is already implicit in (5.16) and (5.25). The financial market equilibrium is unaffected by investors' individual wealth because asset demand is independent of wealth for CARA utility. Information is merely a secondary good that helps investors make better portfolio decisions. So, the demand for information within in the news watcher group is going to be unaffected by wealth as well. Therefore, since investors only differ by level of wealth due to assumptions 5.1, 5.2 and 5.8, whatever is optimal for one group member will be optimal for all other group members. It is thus an admissible simplification to only consider one group representative from now on.

5.4.2 Incentives and externalities

This subsection will take a further step towards deriving the equilibrium at news stands. Without knowing the equilibrium levels of λ^* and N^* yet, we can already establish properties that any information equilibrium must exhibit.

To choose the number of newspapers N, the representative news watcher takes a look at her *ex ante* utility. Similarly, a price watcher looks at his respective *ex ante* utility to see how the signal choice of the news watcher group affects him as an externality. Taking *ex ante* expectations of (5.11), the *ex ante* utility of any investor $i = PW, NW$ is

$$\mathbb{E}^i_{pre}\left[U^{i,*}\right] = -k^i \cdot e^{A\frac{R}{1+R}(F^i + cN^i)} \cdot \mathbb{E}^i_{pre}\left[e^{-\frac{1}{2}\frac{(\tau^i)^2}{1+R}\left(\frac{\mu^i - RP}{(\tau^i)^2}\right)^2}\right] \tag{5.26}$$

for $k^i \equiv \frac{1+R}{R}(\delta R)^{\frac{1}{1+R}} \exp\left\{-A\frac{R}{1+R}W_0^i\right\} > 0$. I have used the short hand $F^i \equiv \mathbf{1}(N^i \geq 1) \cdot F$ in (5.26) again to indicate that only news watchers have to pay the fixed cost if they buy at least one newspaper. Since news watchers are required to buy the same amount of signals N, we can formally also define $N^i \equiv \mathbf{1}(i = NW) \cdot N$ here.

The key term in (5.26) is again

$$\frac{\mu^i - RP}{\tau^i} = \tau^i \, \frac{\mu^i - RP}{(\tau^i)^2}.$$

Given the closed-form financial market equilibrium of lemma 5.3, this term can be expressed in closed form as a function of λ, N, and parameters for all investors $i = PW, NW$. Parameters are: the interest factor R; the prior means and variances $\bar{\mu}_\theta$, $\bar{\tau}^2_\theta$, σ^2_S; $\bar{x}$, ω^2_X; the degree of risk aversion A; the discount factor δ; and the number of investors I (initial wealth W_0^i is irrelevant due to CARA utility). The particular solutions are less important than their properties. So, the explicit terms are not reported here but relegated to appendix C.4 (p. 272). As will become clear shortly, what matters for information acquisition are the two *ex ante* moments of the key term. These two moments are reported in appendix C.5 (p. 272).

We know that the subjective variance of the dividend $(\tau^i)^2$ is certain for all investors (see (5.20) and (5.24)). We also know that both the posterior mean of the dividend μ^i is a sum of normal variables (see (5.17) and (5.21)) and the opportunity cost RP is a sum of normal variables (see (5.16)). Since the sum of normal variables is normally distributed, all investors can apply another convenient fact of the normal distribution—fact C.4 in appendix C.1 (p. 268)—to (5.26) and find their *ex ante* utility to be

$$\begin{aligned} \mathbb{E}^i_{pre}\left[U^{i,*}\right] &= -k^i \cdot \exp\left\{A\frac{R}{1+R}(F^i + cN^i)\right\} \\ &\cdot \frac{1}{\sqrt{1+\frac{(\tau^i)^2}{1+R}\mathbb{V}^i_{pre}\left(\frac{\mu^i - RP}{(\tau^i)^2}\right)}} \exp\left\{-\frac{1}{2}\frac{(\tau^i)^2}{1+R}\frac{\left(\mathbb{E}^i_{pre}\left[\frac{\mu^i-RP}{(\tau^i)^2}\right]\right)^2}{1+\frac{(\tau^i)^2}{1+R}\mathbb{V}^i_{pre}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}\right\}. \end{aligned} \tag{5.27}$$

Since $\mathbb{E}^i_{pre}\left[U^{i,*}\right]$ is negative for CARA utility, any change that brings (5.27) closer to zero is beneficial. Hence, *ex ante* utility is increasing in the *ex ante* expected excess return of the risky asset $\mathbb{E}^i_{pre}\left[\mu^i - RP\right]$, as it should be. However, the variance of the expected excess return $\mathbb{V}^i_{pre}\left(\mu^i - RP\right)$ has an ambiguous effect on *ex ante* utility. On the one hand, an investor finds a higher variance of the expected excess return bad because he is risk averse. On the other hand, he knows that a higher variance of the difference between her return and the market expectations also makes it more likely, on average, that his portfolio yields a lot. So, the double role of the asset price as opportunity cost and information provider also imposes a double role on the variance of $(\mu^i - RP)$.

The representative news watcher maximizes *ex ante* utility $\mathbb{E}^{NW}_{pre}\left[U^{NW,*}\right]$ with respect to the number of newspapers N, given a share λ of news watchers. While making her choice, she does not take into account how N affects *ex ante* utility of the $1-\lambda$ price watchers. Even though the representative news watcher has to choose a discrete number of signals, it is instructive to take the derivative of (5.27) with respect to N. Under certain regularity conditions, the resulting condition could even be interpreted as close to a necessary first order condition for an optimal choice of N

when set to zero.[8] However, I will not use it as a first order condition. Instead, I will use it as a tool to investigate whether the derivative has a certain sign, positive or negative, in general. Then it does not matter whether N is discrete or perfectly divisible. Similarly, the derivative of $\mathbb{E}_{pre}^{PW}\left[U^{PW,*}\right]$ with respect to N can be seen as representing the externality that an additional signal inflicts on price watchers. Taking the derivative and multiplying by the positive factor $-(1+R)/\mathbb{E}_{pre}^{i}\left[U^{i,*}\right]$ yields

$$\begin{aligned}
-\frac{1+R}{\mathbb{E}_{pre}^{i}\left[U^{i,*}\right]}\frac{\partial\mathbb{E}_{pre}^{i}\left[U^{i,*}\right]}{\partial N} \;=\; & - A\,R\,c\;\mathbf{1}(i=NW) \\
& + E^{i}(\lambda,N)\cdot\left[\varepsilon^{i}_{\tau^2,N}(\lambda,N)+\varepsilon^{i}_{\mathbb{E},N}(\lambda,N)\right] \\
& + V^{i}(\lambda,N)\cdot\left[\varepsilon^{i}_{\tau^2,N}(\lambda,N)+\tfrac{1}{2}\varepsilon^{i}_{\mathbb{V},N}(\lambda,N)\right]\cdot\frac{\Delta^{i}(\lambda,N)}{1+R}\,.
\end{aligned} \tag{5.28}$$

The functions $\varepsilon^{i}_{y,N}(\lambda,N)$ denote the elasticity of y with respect to N. For example, $\varepsilon^{i}_{\mathbb{E},N}$ denotes the elasticity of $\mathbb{E}_{pre}^{i}\left[(\mu^i-RP)/(\tau^i)^2\right]$ with respect to N. The definitions of the terms $E^i(\lambda,N), V^i(\lambda,N)$, and $\Delta^i(\lambda,N)$ are

$$E^{i}(\lambda,N) \;\equiv\; \frac{1}{N}\,\frac{\left(\mathbb{E}_{pre}^{i}\left[\frac{\mu^i-RP}{(\tau^i)^2}\right]\right)^2}{\frac{1}{(\tau^i)^2}+\frac{1}{1+R}\mathbb{V}_{pre}^{i}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}, \tag{5.29}$$

$$V^{i}(\lambda,N) \;\equiv\; -\frac{1}{N}\,\frac{\mathbb{V}_{pre}^{i}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}{\frac{1}{(\tau^i)^2}+\frac{1}{1+R}\mathbb{V}_{pre}^{i}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}, \tag{5.30}$$

$$\Delta^{i}(\lambda,N) \;\equiv\; \frac{\left(\mathbb{E}_{pre}^{i}\left[\frac{\mu^i-RP}{(\tau^i)^2}\right]\right)^2-\frac{1+R}{(\tau^i)^2}-\mathbb{V}_{pre}^{i}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}{\frac{1}{(\tau^i)^2}+\frac{1}{1+R}\mathbb{V}_{pre}^{i}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}, \tag{5.31}$$

respectively.

The derivative (5.28) has an intuitive interpretation: The term ARc is the marginal utility loss from an additional signal as it reduces wealth. The second term on the right hand side of (5.28) reflects the marginal utility change that stems from a change in $\tau^i\,\mathbb{E}_{pre}^{i}\left[(\mu^i-RP)/(\tau^i)^2\right] = \mathbb{E}_{pre}^{i}\left[(\mu^i-RP)/\tau^i\right]$. Similarly, the third term reflects the change of the variance and its impact on

[8] If $\mathbb{E}_{pre}^{i}\left[U^{i,*}\right]$ is changing monotonously in N, for example, the resulting condition is fine in the following sense: The condition gives rise to a continuous function $N^*(\cdot)$ of parameters that would indicate optimal signal choices in the continuous case, and, for $N^*\in\mathbb{N}_0$, it coincides with a step function $\hat{N}^*(\cdot)$ that captures the optimal signal choices in the discrete case.

Table 5.1: Signs of elasticities

	$i = PW$	compare	$i = NW$
$\varepsilon_{(\mathbb{E}^i_{pre}[\theta]-RP),N}$ [a]	< 0	$=$	< 0
$\varepsilon^i_{\tau^2,N}$ [b]	< 0	$\varepsilon^{PW}_{\tau^2,N} < \varepsilon^{NW}_{\tau^2,N}$ $\Leftrightarrow \lambda IN > \frac{A\sigma^2_S\ \omega_X}{\bar{\tau}_\theta}$	< 0
$\varepsilon^i_{\tau^2,N} + \varepsilon^i_{\mathbb{E},N}$ [b]	< 0	$=$	< 0
$\varepsilon^i_{\tau^2,N} + \frac{1}{2}\varepsilon^i_{\mathbb{V},N}$ [b]	< 0	$<$	*ambiguous*

[a] This follows from (C.20) with (C.22) and (C.23) with (C.25) in appendix C.5, p. 272.

[b] Elasticities are reported as (C.27) through (C.32) in appendix C.6, p. 274.

utility: $(\tau^i)^2\,\mathbb{V}^i_{pre}\left((\mu^i - RP)/(\tau^i)^2\right) = \mathbb{V}^i_{pre}\left((\mu^i - RP)/\tau^i\right)$. The factor $\Delta^i(\lambda, N)$ is proportional to

$$\left(\mathbb{E}^i_{pre}\left[\frac{\mu^i - RP}{\tau^i}\right]\right)^2 - (1+R) - \mathbb{V}^i_{pre}\left(\frac{\mu^i - RP}{\tau^i}\right)$$

and can thus be positive or negative. It reflects the ambiguous effect that an increase in the variance has on *ex ante* utility.

Table 5.1 displays signs of elasticities. They indicate how more newspapers affect important variables that enter investors' *ex ante* utility. Just as under fully revealing prices before, the *ex ante* expected excess return $(\mathbb{E}^i_{e.a}[\theta] - RP)$ is falling in N for both groups of investors. Price watchers and news watchers even perceive the relative strength of this effect as the same (first row of table 5.1). More information brings expected return and opportunity cost closer to each other while individual beliefs become more similar to market beliefs. This is bad (as row three confirms). However, more information also reduces the dividend's *ex ante* variance for both types of investors (second row). This can be good or bad for utility. Moreover, the portfolio variance, that is the *ex ante* variance of the key term, falls for price watchers, but it may rise of fall for news watchers (last row). To make more definitive statements we need to look at the complete terms in condition (5.28).

Table 5.2 gives an overview of the signs of major terms in condition (5.28). The first row is no surprise any longer: More information has a negative impact on utility because it reduces the

Table 5.2: Signs of utility effects

	$i = PW$	comp.	$i = NW$
$E^i \cdot (\varepsilon^i_{\tau^2,N} + \varepsilon^i_{\mathbb{E},N})$ [a]	< 0	$>$	< 0
$V^i \cdot (\varepsilon^i_{\tau^2,N} + \frac{1}{2}\varepsilon^i_{\mathbb{V},N})$[a]	> 0	-	*ambiguous*
$\Delta^i/(1+R)$[a, b]	$< 0 \;\Leftrightarrow$ $\bar{x}^2 < (\bar{x}_c^{\Delta,PW})^2$	$<$	$< 0 \;\Leftrightarrow$ $\bar{x}^2 < (\bar{x}_c^{\Delta,NW})^2$

[a]For derivations see appendix C.6, p. 273.

[b]Definitions of the threshold values $\bar{x}_c^{\Delta,i}$ are given in (C.39) and (C.40), p. 278.

ex ante expectation of the key term $\mathbb{E}^i_{pre}\left[(\mu^i - RP)/\tau^i\right]$. Both price and news watchers agree that they dislike this. Price watchers perceive this negative effect as less pronounced in absolute terms. They only put a little more weight on the price when extracting information, and a little less weight on their priors. This brings the expected value $\mathbb{E}^{PW}_{pre}\left[\mu^{PW}\right]$ a little closer to the price, but not too much. News watchers, however, do feel the reduction in $\mathbb{E}^{NW}_{pre}\left[(\mu^{NW} - RP)/\tau^{NW}\right]$ from both sides. First, the signal realizations enter $\mathbb{E}^{NW}_{pre}\left[\mu^{NW}\right]$ directly and news watchers start putting more weight on the signal realizations, and less on their priors. Since some other investors also receive the same signals, this brings $\mathbb{E}^{NW}_{pre}\left[\mu^{NW}\right]$ closer to RP. At the same time, price watchers start updating their believes, and the price RP also starts moving closer to $\mathbb{E}^{NW}_{pre}\left[\mu^{NW}\right]$. For news watchers, the excess return is narrowed with double speed, so to say.

Overall, the impact of an additional newspaper on utility is ambiguous for news watchers. The reason is that the effect of an additional signal on the variance $\mathbb{V}^{NW}_{pre}\left((\mu^{NW} - RP)/\tau^{NW}\right)$ is indeterminate (second row in table 5.2). So, there is hope that news watchers are going to acquire information in equilibrium, but they might also prefer no newspaper at all.

Things are more immediate for price watchers. If the stock market is a small market, that is if the expected supply of risky assets is smaller than a cutoff value so that $\bar{x}^2 < \left[\bar{x}_c^{\Delta,PW}(\lambda, N)\right]^2$, then $\Delta^{PW} < 0$. As a consequence, any signal to news watchers must have the character of a pure negative externality for price watchers if it falls below the threshold $\bar{x}_c^{\Delta,PW}(\lambda, N)$ in absolute value (second and third row in table 5.2). What if markets are large in size so that $\Delta^{PW} \geq 0$? Can it happen that this effect becomes so strong that the entire condition (5.28) turns positive for price watchers? As it turns out, the answer is no. The positive effect of more information through the variance can never outweigh the negative effect through a diminished excess return. So, in the present model, more information always inflicts a strictly negative externality on price watchers.

Proposition 5.3 *Any signal to news watchers inflicts a negative externality on price watchers in a two-group rational expectations equilibrium (definition 5.2).*

Proof. In appendix C.7, p. 279. ■

This is a strong result. One might imagine that, when markets are very large in size and the noise in price matters little, price watchers could extract extremely much information from price, and strongly benefit from the variance-lowering effect. This is not the case in the current framework but may again have to do with special properties of the normal-normal conjugate pair. The utility-reducing effect of a shrinking excess return is always stronger.[9]

When taking her decision about newspaper acquisition, the news watcher representative does not care about this externality of her decision. She exclusively considers her private incentives. And her incentives happen to coincide with all other news watchers' incentives because the only difference among them is their initial wealth W_0^i, and that does not matter as condition (5.28) shows. Evaluating condition (5.28) is difficult in general since the effect of an additional newspaper on news watchers' *ex ante* variance $\mathbb{V}_{pre}^{NW}\left((\mu^{NW} - RP)/\tau^{NW}\right)$ is ambiguous (table 5.2). It is therefore instructive to see how condition (5.28) behaves in the limits.

Imagine the extreme case that the club of news watchers has attracted every single investor.

[9] This raises the question whether price watchers should rather stop watching. Would it be rational to stay ignorant? To answer this question, we have to alter the equilibrium concept because news watchers rationally anticipate that price watchers prefer to ignore the information in price. This changes how price responds to more information. The resulting equilibrium remains to be investigated in future research.

Then the incentive to acquire newspapers becomes

$$\lim_{\lambda\to 1} -\frac{1+R}{\mathbb{E}_{pre}^{NW}\left[U^{NW,*}\right]}\frac{\partial \mathbb{E}_{pre}^{NW}\left[U^{NW,*}\right]}{\partial N} \quad = \quad -A\,R\,c\, - \frac{(1+R)A^2\sigma_S^2\bar{\tau}_\theta^4}{\sigma_S^2+N\bar{\tau}_\theta^2} \cdot \frac{I^2(1+R)(\sigma_S^2+N\bar{\tau}_\theta^4)(\bar{x}^2+\omega_X^2)+A^2\sigma_S^2\bar{\tau}_\theta^2\omega_X^4}{\left[I^2(1+R)(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\bar{\tau}_\theta^2\omega_X^2\right]^2} < 0. \tag{5.32}$$

In other words, there is a strong disincentive to receive information even if newspapers are for free. This establishes

Proposition 5.4 *There must be at least one price watcher in a two-group RICE (definition 5.2).*

In any rational expectations equilibrium (definition 5.1),

- *either at least one investor must receive less signals than any other investor (when signals are sold in perfect copies),*
- *or at least one investor must acquire a signal that no other investor has received.*

Proof. Suppose every investor became a news watcher in a two-group rational expectations equilibrium (definition 5.2), then news watchers would want to acquire no signal by (5.32), a contradiction. The limit (5.32) itself follows from (C.34), (C.36), and (C.38) in appendix C.6 (p. 275).

Now consider the more general case of definition 5.1. If signals are sold in perfect copies as supposed in this section, then a symmetric equilibrium under definition 5.1 in which all investors receive the same number of signals coincides with an equilibrium that involves $\lambda = 1$ under definition 5.2. However, this kind of equilibrium does not exist. So, at least one investor must receive less of the same signals or at least one investor must receive a signal that nobody else received in any rational expectations equilibrium. ∎

In other words, a symmetric equilibrium cannot exist in which all investors would receive a positive number of copies of the same newspapers. It doesn't even exist if newspapers are free of charge. The statement shows that the present definition of a two-group equilibrium is not so restrictive after all. There must be at least two groups of differently informed investors in any rational expectations equilibrium. They may not choose to receive the same copies, though. More generally, proposition 5.4 further supports the insight that investors dislike agreement and do not want information to become too common.

The most important question is, however, what happens in the other extreme. What does condition (5.28) look like when there is no news watcher yet? The potential news watcher representative asks herself whether she should start a news watcher club with one member—herself. As it turns out,

$$\begin{aligned}\lim_{\lambda\to 0} -\frac{1+R}{\mathbb{E}^{NW}_{pre}\left[U^{NW,*}\right]}\frac{\partial \mathbb{E}^{NW}_{pre}\left[U^{NW,*}\right]}{\partial N} \quad &= \quad -A\,R\,c \\ &+\frac{I^2(1+R)\sigma_S^2\bar{\tau}_\theta^2}{2(\sigma_S^2+N\bar{\tau}_\theta^2)\left[I^2\left[(1+R)\sigma_S^2+N\bar{\tau}_\theta^2\right]+A^2\bar{\tau}_\theta^2\omega_X^2(\sigma_S^2+N\bar{\tau}_\theta^2)\right]^2} \\ &\cdot\left(I^2\left[(1+R)\sigma_S^2+N\bar{\tau}_\theta^2\right]+A^2\bar{\tau}_\theta^2(\sigma_S^2+N\bar{\tau}_\theta^2)(\omega_X^2-\bar{x}^2)\right).\end{aligned} \tag{5.33}$$

Even if she were to become the solely informed investor in the market, a representative news watcher may prefer to remain dumb. Apart from the uninteresting case of a prohibitively high newspaper price c, this is likely to occur if $\bar{x}^2$ is relatively high compared to ω_X^2. Then the last factor in (5.33) can turn negative. In other words, nobody may want to become informed, if markets are large! Why? In large markets, given a level of noise in the price ω_X^2, the asset prices is very informative for price watchers. News watchers know that price watchers will start putting a lot of weight on the observed price and little weight on their priors. As a result, news watchers must rationally anticipate that the expected excess return $\mathbb{E}^{NW}_{pre}\left[\mu^{NW}-RP\right]$ is falling quite strongly with every newspaper as opportunity cost RP moves closer to $\mathbb{E}^{NW}_{pre}\left[\mu^{NW}\right]$ while price watchers are updating their beliefs. So, market size and informativeness of the price are closely linked for the incentives to acquire information.

Formally, condition (5.33) implies that, in the limit where $\lambda = 0$, the threshold of market size is given by

$$\lim_{\lambda\to 0} -\frac{1+R}{\mathbb{E}^{NW}_{pre}\left[U^{NW,*}\right]}\frac{\partial \mathbb{E}^{NW}_{pre}\left[U^{NW,*}\right]}{\partial N} < 0 \quad \Leftrightarrow \quad \bar{x}^2 > \left[\bar{x}^{news}_{c,\lambda=0}(N;c)\right]^2,$$

with a cutoff value

$$\bar{x}^{news}_{c,\lambda=0}(N;c)^2 \equiv \frac{I^2\left((1+R)\sigma_S^2 N\bar{\tau}_\theta^2\right) + A^2\bar{\tau}_\theta^2\omega_X^2(\sigma_S^2+N\bar{\tau}_\theta^2)}{I^2A^2\bar{\tau}_\theta^2(\sigma_S^2+N\bar{\tau}_\theta^2)} \cdot\left(1 - \frac{2RA(\sigma_S^2+N\bar{\tau}_\theta^2)\left(I^2\left((1+R)\sigma_S^2N\bar{\tau}_\theta^2\right)+A^2\bar{\tau}_\theta^2\omega_X^2(\sigma_S^2+N\bar{\tau}_\theta^2)\right)}{I^2(1+R)\sigma_S^2\bar{\tau}_\theta^2}\,c\right). \tag{5.34}$$

This threshold level $\bar{x}^{news}_{c,\lambda=0}(N;c)^2$ is strictly falling in N. So, the incentive for at least one investor to become a news watcher is the stronger the lower the prospective number of newspapers.

Under the assumption that condition (5.28) is maximal at $\lambda = 0$, the following claim can be made.

Claim 5.1 *Suppose the incentive to acquire a signal is strongest when $\lambda = 0$, then the following is true.*

There are only price watchers in a two-group rational expectations equilibrium (definition 5.2) if risky asset supply exceeds a threshold such that $\bar{x}^2 \geq \left[\bar{x}^{news}_{c,\lambda=0,N=0}(c)\right]^2$.

Proof. Under the assumption made, condition (5.28) takes its maximum at $\lambda = 0$ (on the interval $\lambda \in [0,1]$). So, if (5.28) is smaller than zero at $\lambda = 0$, it cannot exceed zero for any other value of λ, given N. Thus, no news watcher would want to buy a signal under a sufficient condition that forces (5.33) below zero, and there will only be price watchers in a two-group rational expectations equilibrium (definition 5.2).

The limit (5.33) follows from (C.34), (C.36), and (C.38) in appendix C.6 (p. 275). It is linear in $\bar{x}^2$. Setting it equal to zero, and solving out for $\bar{x}^2$ yields the threshold level (5.34). Condition (5.34) is sufficient for no signal acquisition to occur under the assumption made. It is not a necessary condition due to the indivisibility of signals. ∎

Since investors can go long or short in the risky asset, this result depends on market size in absolute value (the square of $\bar{x}$). Information is the more valuable for news watchers the smaller markets are. The reason is that market size is just the flip side of the informativeness of price. News watchers do not want price watchers to free-ride on their newspaper acquisitions because that

Table 5.3: Incentives for news watchers in the limit

	$E^{NW}(\varepsilon^{NW}_{\tau^2,N} + \varepsilon^{NW}_{\mathbb{E},N})$	$V^{NW}(\varepsilon^{NW}_{\tau^2,N} + \frac{1}{2}\varepsilon^{NW}_{\mathbb{V},N})\frac{\Delta^{NW}}{1+R}$
$\lim_{\omega_X \to 0}$[a]	$-\frac{A^2\sigma_S^2\bar{\tau}_\theta^4}{I^2(\sigma_S^2+N\bar{\tau}_\theta^2)^2}\bar{x}^2$	0
$\lim_{\omega_X \to \infty}$[a]	0	$-\frac{(1+R)\lambda\bar{\tau}_\theta^2}{\sigma_S^2+\lambda N\bar{\tau}_\theta^2}$
$\lim_{A \to 0}$[a]	0	0
$\lim_{A \to \infty}$[a]	$-\frac{(1+R)\lambda\bar{\tau}_\theta^2}{(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\omega_X^2}\bar{x}^2$	$\frac{(1+R)\lambda\bar{\tau}_\theta^2}{(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\omega_X^2}(\bar{x}^2-\omega_X^2)$
$\lim_{1/\sigma_S \to 0}$[a]	0	0
$\lim_{1/\sigma_S \to \infty}$[a]	0	0

[a] Limits follow from (C.34) and (C.36) in appendix C.6, p. 275.

reduces the expected excess return of the risky asset. So, the larger markets, the more informative prices are for price watchers, and the stronger the negative effect of price watchers' updating on news watchers utility. The close relation between market size and the informativeness of price would not change if the noise in the price system came from another source than asset supply. Looking (far back) at the structure of equilibrium price in (5.16), we could also have added Gaussian white noise to the price, and results would have carried over.

Table 5.3 reports some further noteworthy limits. When the price system becomes extremely informative as $\omega_X \to 0$, news watchers perceive the negative impact on the excess return more strongly than the positive impact on the variance of the excess return and prefer to be price watchers. This is nothing but the extreme case of the preceding section 5.3 where prices were fully revealing. On the other extreme, when the price system ceases to be informative as $\omega_X \to \infty$, news watchers must not fear a negative impact in excess return any more. However, the potentially positive effect on the variance of the excess return turns negative because, if ω_X^2 is large, RP gets more noisy with more newspapers. When investors become extremely risk averse ($A \to \infty$), they lose their interest in risky assets and consequently their interest in information. In all these cases, news watchers would not even want to accept a signal for free.

When investors become risk neutral ($A \to 0$), they do not mind receiving signals for free, but they would never want to pay for it—as lemma 5.1 establishes in general. Similarly, when a signal is absolutely imprecise ($1/\sigma_S^2 \to 0$) investors are indifferent about receiving it or not: It does neither harm nor good, but never pay for it. Finally, when signals become absolutely precise and reveal the realization of θ itself as $1/\sigma_S^2 \to \infty$, news watchers would accept it but not pay for it. Such an infinitely precise signal turns the previously risky asset into a second, riskless bond and mandates that RP equal θ/R. In this extreme case, the two assets become perfect substitutes.

5.4.3 The information market equilibrium

The previous subsection characterized properties of an equilibrium at the news stands. It remains to derive this information equilibrium itself. Suppose again that we can treat the number of signals N as if it were close to perfectly divisible.[10] Then, the news watcher representative can decide the equilibrium amount of information by looking at condition (5.28) and setting it to zero

$$\begin{aligned} -\frac{1+R}{\mathbb{E}_{pre}^{NW}\left[U^{NW,*}\right]}\frac{\partial \mathbb{E}_{pre}^{NW}\left[U^{NW,*}\right]}{\partial N} &= -\,A\,R\,c \\ &+\frac{1}{N}\;E^{NW}(\lambda,N)\left[\varepsilon_{\tau^2,N}^{NW}(\lambda,N)+\varepsilon_{\mathbb{E},N}^{NW}(\lambda,N)\right] \\ &+\frac{1}{N}\;V^{NW}(\lambda,N)\left[\varepsilon_{\tau^2,N}^{NW}(\lambda,N)+\tfrac{1}{2}\varepsilon_{\mathbb{V},N}^{NW}(\lambda,N)\right]\frac{\Delta^{NW}(\lambda,N)}{1+R}=0. \end{aligned} \tag{5.35}$$

The news watcher representative chooses N given the share λ of members in the news watcher club (definition 5.2). So, the above condition implies an equilibrium amount of signals $N^*(\lambda;c)$. Unfortunately, the acquisition rule $N^*(\lambda;c)$ has no closed form (but can be shown to be a polynomial of ninth degree). Things are getting easier, however, if we look at them graphically.

The falling curve in figure 5.2 is a plot of condition (5.35).[11] It shows combinations of N and λ for which (5.35) is satisfied. Or, put in economic terms, this curve shows the optimal choice of N^* if it were continuous. The curve shifts to the Southwest when the cost of a signal c increases as can be seen from (5.35). Since signal choice has to be concrete, however, the optimal choice of N^*

[10] See footnote 8 (p. 164).

[11] The underlying parameter values are $I = 100$; $c = .005$, $F = 10c$; $R = 1.1$, $\bar{\mu}_\theta = 1.3$; $\bar{x} = 1$; $A = 1$, $\sigma_S = 1$, $\bar{\tau}_\theta = 1$, $\omega_X = 100$; $\delta = .9$, $W = 1$.

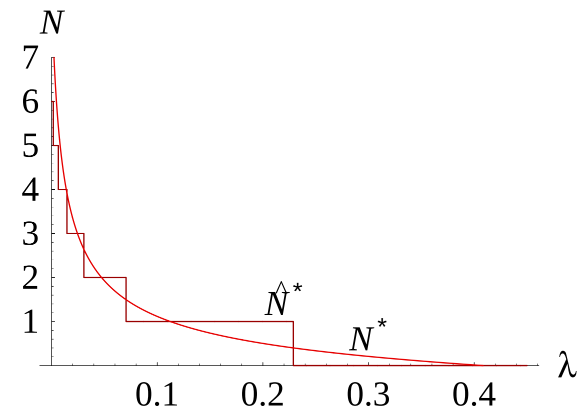


Figure 5.2: Optimal choice of the number of signals

given λ is a step function $\hat{N}^*(\lambda; c)$. Figure 5.2 also depicts this proper *newspaper acquisition curve.* The two curves show that condition (5.35) does a pretty good job for a large number of signals, but is not so helpful when N gets small.

Any equilibrium must occur along the newspaper acquisition curve $\hat{N}^*(\lambda; c)$. Given news watchers' anticipated choice of $\hat{N}^*(\lambda; c)$, each investor decides whether to become a price watcher or a news watcher. In equilibrium, every news watcher must find it preferable to remain a news watcher. Her *ex ante* utility must be weakly higher than a price watcher's *ex ante* utility. Formally, $\mathbb{E}_{pre}^{NW}\left[U^{NW,*}(N,\lambda;c,F)\right] \geq \mathbb{E}_{pre}^{PW}\left[U^{PW,*}(N,\lambda)\right]$. Similarly, every price watcher must find it preferable not to become a news watcher. This implies $\mathbb{E}_{pre}^{PW}\left[U^{PW,*}(N,\lambda)\right] \geq \mathbb{E}_{pre}^{NW}\left[U^{NW,*}(N,\lambda;c,F)\right]$. As a result,

$$\mathbb{E}_{pre}^{NW}\left[U^{NW,*}(N,\lambda;c,F)\right] - \mathbb{E}_{pre}^{PW}\left[U^{PW,*}(N,\lambda)\right] = 0 \tag{5.36}$$

must hold in equilibrium. Given news watchers' signal choice N, this condition implies an equilibrium share of news watchers $\lambda^*(N; c, F)$. It also implies that the initial wealth of investors within each group must be the same if the same fixed information cost F applies to everyone. To keep things interesting, make a final

Assumption 5.13 (Same wealth) *Initial wealth is identical, $W_0^i = W_0$, across all investors $i =$*

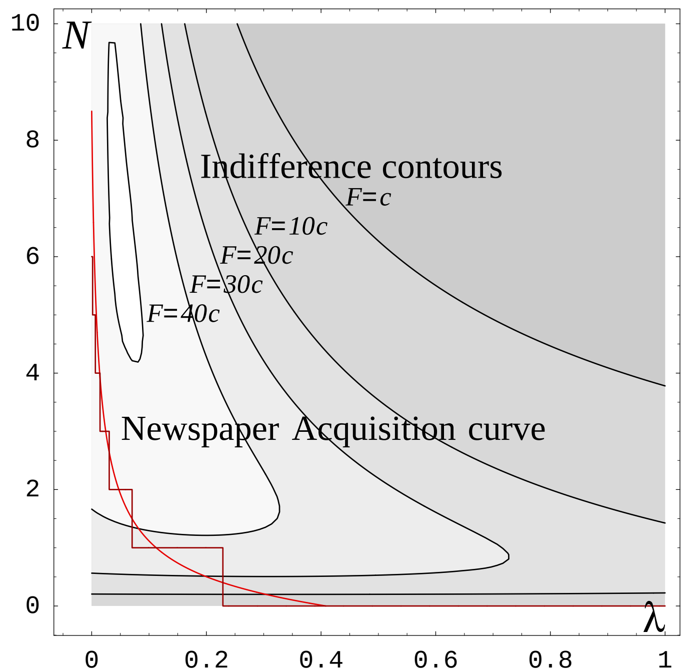


Figure 5.3: Equilibrium combinations of the number of signals and the share of newswatchers

$1, ..., I$.

This assumption would not be needed if we allowed for more than only two groups of investors. Then, however, no closed-form financial market equilibrium would exist.

In equilibrium, every news watcher receives N^* signals for c dollars each and pays the fixed cost F. So, a news watcher's *ex ante* utility depends on both c and F. Thus, condition (5.36) also depends on c and F implicitly. As a consequence, it is not of much concern that both N and $\lambda \equiv I^{NW}/I$ are not perfectly divisible. Either the newspaper price c or the fixed information cost F, or both, adjust to clear the markets accordingly.

Figure 5.3 shows contour plots of condition (5.36) for various levels of the fixed cost F.[12] These indifference contours need not satisfy a functional relationship between N and λ. In fact, they are mostly correspondences. By varying F we can find a combination of N and λ that lies on the

[12]Parameter values are the same as in figure 5.2. In addition, $W = 1$. See footnote 11 (p. 172).

newspaper acquisition curve *and* on an according indifference contour. This is one equilibrium. By varying F further, we can find several additional combinations of N and λ that lie on the newspaper acquisition curve at some other point. So, the information equilibrium need not be unique. For many different levels of F one may find an equilibrium pair $(N^*(c, F), \lambda^*(c, F))$ that makes this particular F an equilibrium price together with some unit price c that is implicit in both the newspaper acquisition curve and the indifference contour.

Lemma 5.4 *Countably many two-group rational expectations equilibria (definition 5.2) may exist.*

Proof. The number of equilibria must be countable because N is an integer. Parameters permitting, we can construct examples as in figure 5.3 in which multiple equilibria can be found by varying the fixed information cost F. ■

So, the equilibrium at the news stands at 9am need not be unique, whereas the partial equilibrium at Wall Street at 10am will be unique *given* N^* and λ^*. The number of signals N has to be discrete. This makes it hard to derive general conditions under which there are at least two equilibria.

The previous argument also suggests that the fixed information cost F will take a strictly positive value in many equilibria. In fact, it must always be strictly positive. Recall that price watchers suffer a negative externality and are strictly worse off than news watchers if the fixed information cost F is zero. Consequently, no information equilibrium with a positive amount of information can exist for $F = 0$ as long as at least one investor has an incentive to become a news watcher.

Proposition 5.5 *An equilibrium (definition 5.2) with at least one news watcher requires a fixed information cost F that is strictly positive.*

Proof. Suppose there is at least one news watcher, then $N^* \geq 1$. In addition, by proposition 5.4 there must be at least one price watcher, $\lambda^* < 1$. Further suppose that $F = 0$. Since a news watcher is free to choose N, it must be the case that $\mathbb{E}_{pre}^{NW}\left[U^{NW,*}(N^*)\right] \geq \mathbb{E}_{pre}^{NW}\left[U^{NW,*}(N=0)\right]$ by revealed preference. By proposition 5.3 price watchers face a negative externality so that they suffer a utility loss $\mathbb{E}_{pre}^{PW}\left[U^{PW,*}(N \geq 1)\right] < \mathbb{E}_{pre}^{PW}\left[U^{PW,*}(N=0)\right]$. Since, $\mathbb{E}_{pre}^{NW}\left[U^{NW,*}(N=0)\right] = \mathbb{E}_{pre}^{PW}\left[U^{PW,*}(N=0)\right]$ we can infer that $\mathbb{E}_{pre}^{NW}\left[U^{NW,*}(N \geq 1)\right] > \mathbb{E}_{pre}^{PW}\left[U^{PW,*}(N \geq 1)\right]$ for small markets. So, in equilibrium a strictly positive fixed information cost F must bring news watchers' utility down to price watchers' utility. ∎

In the present framework, information has to be priced with a two-part tariff. Otherwise no equilibrium at the news stands would exist as long as at least one investor has an incentive to become a news watcher. Since we know from proposition 5.4 that there must be at least two groups of differently informed individuals in general (equilibrium definition 5.1), the present proposition 5.5 also hints at the general case. If information inflicts a negative externality on at least one investor, at least the best informed group of investors must pay a strictly positive fixed cost to access information. Otherwise no equilibrium exists. Even large markets may require that the fixed information cost is strictly positive in equilibrium, but they need not. In general, the utility difference (5.36) is a complicated function of λ, N, c, and F.[13]

5.4.4 Informational efficiency

Are the equilibria at news stands informationally efficient? That is, would a benevolent social planner allocate signals to investors in the same manner? A benevolent social planner maximizes $\sum_{i=1}^{I} \mathbb{E}_{pre}^{i}\left[U^{i,**}\right]$ with respect to $\{N^1, ..., N^I\}$. Since there is no closed-form solution to the financial market equilibrium in general, it is difficult to characterize the unconstrained social optimum. However, we can investigate the welfare properties of two-group equilibria (definition 5.2) in the current context. A benevolent social planner can dictate the news watcher group to buy N^{**} signals for each member, charging every news watcher the marginal cost c of signal provision. To keep

[13] The present model still contains Grossman and Stiglitz' (1980) version as a special case. There are some noteworthy differences in the equilibria that result, however, on which I comment in appendix C.8 (p. C.8).

matters simple, suppose c is precisely the marginal cost of the newspaper copy and does not include the production of the newspaper content, for instance. Then a social planner will find a charge of c for each copy the right price, and we can focus on further welfare aspects of the equilibrium.

So, the social planner will maximize $(1-\lambda)\mathbb{E}^{PW}_{pre}\left[U^{PW,**}\right] + \lambda\mathbb{E}^{NW}_{pre}\left[U^{NW,**}\right]$ given c, where $U^{i,**}$ denotes posterior indirect utility after the social planner interfered at the news stands.[14] Treating signals as if they were perfectly divisible, we can differentiate this weighted average with respect to N, given λ, and find

$$\begin{aligned}
-\frac{1+R}{\mathbb{E}^{NW}_{pre}\left[U^{NW,**}\right]}\,\frac{\partial}{\partial N}&\left((1-\lambda)\mathbb{E}^{PW}_{pre}\left[U^{PW,**}\right] + \lambda\mathbb{E}^{NW}_{pre}\left[U^{NW,**}\right]\right) = \\
&= -\,\lambda\, A\, R\, c \\
&\quad + \lambda\left(E^{NW}\left[\varepsilon^{NW}_{\tau^2,N} + \varepsilon^{NW}_{\mathbb{E},N}\right] + V^{NW}\left[\varepsilon^{NW}_{\tau^2,N} + \tfrac{1}{2}\varepsilon^{NW}_{\mathbb{V},N}\right]\tfrac{\Delta^{NW}}{1+R}\right) \\
&\quad + (1-\lambda)\,\frac{\mathbb{E}^{PW}_{pre}\left[U^{PW,**}\right]}{\mathbb{E}^{NW}_{pre}\left[U^{NW,**}\right]} \\
&\qquad \cdot\left(E^{PW}\left[\varepsilon^{PW}_{\tau^2,N} + \varepsilon^{PW}_{\mathbb{E},N}\right] + V^{PW}\left[\varepsilon^{PW}_{\tau^2,N} + \tfrac{1}{2}\varepsilon^{PW}_{\mathbb{V},N}\right]\tfrac{\Delta^{PW}}{1+R}\right).
\end{aligned} \tag{5.37}$$

From the preceding analysis we know that, in a market equilibrium, at least one investor must be a price watcher. A social planner clearly agrees. For $\lambda = 1$, the last term in condition (5.37) vanishes. Moreover, the second term on the right hand side turns negative: If the news watcher group included every single investor then any newspaper would strictly reduce news watchers' *ex ante* utility (proposition 5.4). So, it cannot be socially desirable that all investors read the same N newspapers, even if newspapers are for free.

We know that any signal inflicts a negative externality on price watchers (from proposition 5.3). So, the last term in condition (5.37) is always negative. In addition, we know that when markets are large in size, news watchers do not even have an incentive to buy a newspaper if the group has only one member (claim 5.1). Again, a social planner agrees.

Proposition 5.6 *An informationally efficient allocation of signals*

[14] The social planner cannot transfer knowledge between investors so that *ex ante* utility is taken with respect to investors' individual *ex ante* beliefs.

- *has to be asymmetric so that at least one investor receives either less or different signals if the allocation involves a positive number of signals;*

- *results in no information acquisition if risky asset supply exceeds a threshold such that* $\bar{x}^2 \geq \left[\bar{x}^{news}_{c,\lambda=0,N=0}(c)\right]^2$ *(provided the incentive to acquire a signal is strongest when* $\lambda = 0$*).*

Proof. By an extension of propositions 5.4, 5.3 and claim 5.1, and condition (5.37). The threshold for market size $\left[\bar{x}^{news}_{c,\lambda=0,N=0}(c)\right]^2$ is given in (5.34). ■

Loosely speaking, a social planner agrees with the market outcomes in the extremes. However, this is not so in general. Since every signal causes a negative externality to price watchers, a social planner would tell news watchers to acquire less signals for every given share λ of news watchers. Suppose signals were perfectly divisible. Then the social planner strictly prefers a signal allocation in which less signals than in the market equilibrium are given to news watchers. Since signals have to be purchased in integer numbers, however, the social planner might settle with the market outcome for ranges of equilibria. Market outcomes could be informationally efficient by coincidence, so to say.

Claim 5.2 *When there is at least one news watcher in equilibrium (definition 5.2), markets provide inefficiently much information in the following sense. A benevolent social planner would, for any given* λ^**, allocate strictly less signals if signals were perfectly divisible.*

Proof. By proposition 5.3 and (5.37). ■

Figure 5.4 depicts the newspaper acquisition curves for a news watcher representative and for a social planner under the same parameter as used in the preceding figures.[15] The social planner would rather prefer to implement no information at all in this example, instead of having a group of news watchers read perfect copies of the same newspapers. It remains to be analyzed how general this finding is. An implication may be: A positive amount of information can only

[15] See footnote 11 (p. 172).

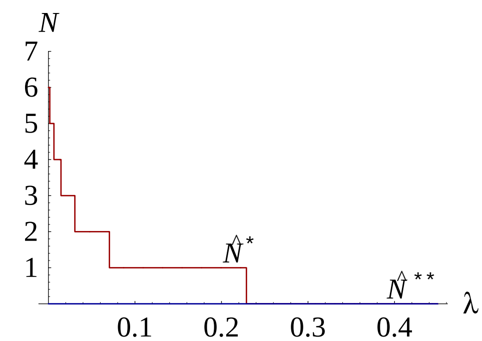


Figure 5.4: Informationally efficient choice of the number of signals

be informationally efficient if news watchers receive different information. A social planner likes investors to hold heterogeneous beliefs. We can conclude, however, that there are market outcomes in two-group rational expectations equilibria that tend to be informationally inefficient. Too many signals are purchased in those equilibria rather than too few as under fully revealing prices and a gamma-Poisson pair of distributions.

5.5 Conclusion

How much information do investors want to hold, and how much should they acquire? To address this question, this chapter 5 considers a rational information choice equilibrium in which both asset markets at Wall Street and newspaper markets clear. Information is not a good or bad in its own right. It is only valuable inasmuch investors anticipate to act upon it. Therefore, risk neutral investors never want to buy information. They only care about the first moment of the asset return and more information cannot change that moment *ex ante* by the law of iterated expectations.

When investors are risk averse, the effect of information on asset price and utility is considered both in equilibria with fully and with partly revealing prices. Gaussian random variables and CARA utility are used to obtain closed-form solutions where possible. Beyond previous work, this model allows investors to choose information sources that may coincide with other investors' sources. In other words, newspapers rather than private detective's reports can be considered.

Grossman and Stiglitz' (1980) paradox that no equilibrium exists if the asset price is fully revealing is resolved in the common framework of the preceding chapter 4 and the present chapter. Under a gamma distributed return, investors bought a positive amount of information whenever markets were large enough, when investors were sufficiently risk averse, or when the variance of the risky asset was relatively high compared to its payoff. Under a normal-normal conjugate pair of distributions, the unique equilibrium involves no information acquisition at all and is socially efficient. The negative effect of commonly available information on relative asset returns makes it undesirable to acquire information under the very specific characteristics of Gaussian random variables. Beliefs become more homogeneous across agents when information becomes common, and the asset price moves closer to every individual investor's return expectations. This diminishes the value of the risky asset from the point of view of each individual investor and outweighs positive effects of information. Under a normally distributed asset return, the negative effect always prevails and no information is acquired. In the case of a gamma distributed asset returns, this happened only if few risky assets were supplied, or if their returns were not very volatile. Otherwise, information had a well-defined and strictly positive value. The main difference between the gamma-Poisson and the normal-normal pair of distributions is that the variance of the asset return is *deterministic* in the normal case and thus independent of the realization of the asset return.

When prices are noisy and only partly informative about other investors' information, the negative effect of common information remains present but is mitigated. Even under the normal-normal pair of distributions, investors start to acquire information as long as markets are sufficiently small so that prices reveal little information to others. When informed investors (news watchers) acquire information, they inflict a negative externality on less informed investors (price watchers) who do not purchase own information but merely observe the price realization. The reason is that price watchers rationally anticipate the arrival of information in the market, simultaneously update their beliefs in the same direction and thus make asset price move closer to their own (and average) beliefs. Therefore, the beneficial effect of more precise information never outweighs the loss from a reduced expected return for the price watchers.

When markets are not too large, so that prices do not become too informative, there is likely a group of investors who prefer to buy information even under the normal-normal conjugate pair—as long as price is not fully revealing. It can never be the case that this group includes all investors if signals are sold in perfect copies. Yet, some fraction of investors may choose to become informed. More information lowers the expected variance of their portfolio, which raises their *ex ante* utility because they are risk averse. This analysis of partly revealing prices is exploratory. Generalizations are needed and the existence of equilibria is a mathematical problem to resolve in the absence of closed-form solutions.

A benevolent social planner agrees that markets should never make everybody equally well informed. Whenever there are some informed investors in equilibrium, markets can become informationally inefficient under certain circumstances because they involve *more* information acquisition than a social planner would implement. Informed investors do not account for the negative externality that they inflict on the less informed. The presence of a negative externality raises the question whether less informed investors should stop extracting information from price and rely on their own priors only. This concern remains a question for future investigation.

What do these findings imply for financial crises? First and foremost, investors do have incentives to conceal information. In the purely rational models of chapters 4 and 5 this is reflected in the fact that investors may refuse to receive information. In extensions of the models with more than one period, investors may thus indeed have reason to hesitate and not act on information as long as they can expect others to hold on to their portfolio positions. By acting, an investor transmits his information to every other investor in the market and works to reduce the value of the asset for himself because everyone's beliefs align with his. On the other hand, not acting on information despite the expected possibility of lower returns worsens the portfolio value. So, in the presence of a possible financial crisis a tradeoff between acting or not acting on information may arise. A precise theoretical analysis remains to be conducted in a model with more than two periods.

This reasoning suggests that a devaluation or revaluation of assets may be delayed because investors who possess superior information are hesitant to change their portfolio. They are reluctant to transmit their information to the market. A delayed response, however, is likely to be stronger and may give rise to the pattern of a crisis. It may therefore be a recommendable policy for a government agency or a central bank to widely publicize any price relevant information that may otherwise not reach a number of investors in the market. However, a theoretical model geared to the specific setting of a financial crisis and its stages is yet to be analyzed.

Chapter 6

An Outlook on Future Research into Globalization

How can economies benefit from globalization? This basic concern motivates the present dissertation. In particular, how can less developed countries engage global markets on their own terms? I address several aspects of this issue. First, I investigate theoretically under what circumstances trade helps reduce the productivity gap between less and more advanced regions. Second, I demonstrate empirically how Brazil's manufacturing firms respond to falling trade barriers. Third, I examine information in financial markets, which, I argue, may determine the timing and prevention of financial crises. Yet, many aspects of these concerns deserve revisiting and open questions await their resolution. Both theoretical and empirical work is called for.

I show in my theoretical chapter on trade and growth that, even in the presence of dynamic externalities, a less developed region can achieve faster growth if it pursues intraindustry trade. The model assumes that productivity gains are largely due to learning by doing. Workers who change employment spread this knowledge throughout a region. Does this frequently made assumption hold up to empirical evidence? Would alternative endogenous growth engines result in similarly favorable findings for North-South trade? Knowledge increasingly drives today's economies. What are the principal engines of knowledge creation? How does globalization affect knowledge creation and dissemination? How can we ensure that many people can benefit from innovations and technical change? What factors, more generally, determine technology adoption and to what degree does technical change depend on existing skills in the labor force? There is a need for both theoretical models that capture firms' choices in equilibrium and for empirical investigations.

I investigate what strategies Brazilian manufacturing firms adopt in response to the changing openness of the Brazilian economy. In an effort to separate effects causally, I find that the use of foreign inputs plays a negligible role in productivity change in the short term. In contrast, foreign competition pushes firms to raise efficiency markedly and forces the least efficient firms out of business. Counterfactual simulations show that especially the competitive push on surviving firms is a salient source of immediate change during and after Brazil's trade reform. However, who benefits? Import-competing sectors shrink in size and shed labor. How do wages change? Exploratory research shows that the real wage rises for workers across all levels of education and tenure in Brazil during and after trade reform. In this sense, it is likely that workers benefit from technical change in absolute terms. Simultaneously, however, the relative wages for workers with different levels of schooling change. Middle schooling levels lose *vis à vis* lower education levels and college graduates gain relative to lower education levels in Brazil during and after trade reform. Are trade liberalization and subsequent technical change causal forces behind these changes in the wage distribution? Future research with data sources at the level of workers, matched to firms, may help address these questions and many related issues. So far, three-dimensional panels that observe firms and their individual workers over time are only available for a select group of countries. However, existing but separate data sources for firms and workers seem to allow the construction of such panels for several economies, among them Brazil. Does knowledge spread when workers change employment? How exactly do firms pursue technical change and what role does the skill-composition of the labor force play? When firms lay off workers to raise efficiency and compete, does human capital get lost in this displacement process? What future wage rates do displaced workers face and where do they move?

Financial crises plagued capital markets in many developing and emerging economies over the past decade. While some countries faced home made problems, others suffered financial shocks that neither their policies nor the fundamentals of their securities seemed to warrant. Information in financial markets may explain a part of the puzzling timing of crises. I propose a theoretical framework in this dissertation that integrates the information choice into a comprehensive rational expectations equilibrium. I confirm under varying theoretical assumptions that investors have incentives to conceal information and may prefer not to act upon it under certain circumstances. Are the results robust in models with different distributional assumptions? Can they be generalized to a family of conjugate prior distributions? The theoretical framework proposed here assumes that investors are price takers at Wall Street but act strategically in the market for information. How does the demand for and the value of information change when investors act strategically in their portfolio choice, too? The framework presents new implications for investor behavior. Do these predictions hold up to empirical confirmation? The Asian financial crisis or speculation against the Australian dollar may be useful settings to test some of the predictions. Finally, how will the findings play out in a multi-period model that is specifically geared to the stages of a financial crisis? What does an applied model of this kind imply for the timing of crises and their prevention?

The four principal chapters of this dissertation aim to provide insights into international trade and finance. They consider prominent concerns of globalization and are conceived to be of foremost relevance to developing economies. It is my hope and aspiration, however, that the essays will be of broad interest for a research agenda on globalization.

Bibliography

Admati, Anat R. and Paul Pfleiderer, "Selling and Trading on Information in Financial Markets," *American Economic Review*, May 1988, *78* (2), 96–103.

Aghion, Philippe and Peter Howitt, *Endogenous Growth Theory*, Cambridge, Massachusetts: MIT Press, 1998.

Allen, Beth E., "Generic Existence of Completely Revealing Equilibria for Economies with Uncertainty when Prices Convey Information," *Econometrica*, September 1981, *49* (5), 1173–99.

Amann, Edmund, "Technological Self-Reliance in Brazil: Achievements and Prospects. Some Evidence from the Non-serial Capital Goods Sector," *Oxford Development Studies*, October 1999, *27* (3), 329–57.

Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *Review of Economic Studies*, June 1962, *29* (3), 155–73.

Avery, Christopher and Peter Zemsky, "Multidimensional Uncertainty and Herd Behavior in Financial Markets," *American Economic Review*, September 1998, *88* (4), 724–48.

Aw, Bee Yan, Sukkyun Chung, and Mark J. Roberts, "Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China)," *World Bank Economic Review*, January 2000, *14* (1), 65–90.

Back, Kerry, "Insider Trading in Continuous Time," *Review of Financial Studies*, 1992, *5* (3), 387–409.

——**, C. Henry Cao, and Gregory A. Willard**, "Imperfect Competition among Informed Traders," *Journal of Finance*, October 2000, *55* (5), 2117–55.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, August 1992, *107* (3), 797–817.

Bardhan, Pranab K., "The Contributions of Endogenous Growth Theory to the Analysis of Development Problems: An Assessment," in Jere R. Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*, Vol. 3, New York: North-Holland, 1995, chapter 46, pp. 2983–98.

Barlevy, Gadi and Pietro Veronesi, "Information Acquisition in Financial Markets," *Review of Economic Studies*, January 2000, *67* (1), 79–90.

Bartelsman, Eric J. and Mark Doms, "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, September 2000, *38* (3), 569–94.

Ben-David, Dan, "Equalizing Exchange: Trade Liberalization and Income Convergence," *Quarterly Journal of Economics*, 1993, *108* (3), 653–79.

Benabou, Roland and Guy Laroque, "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics*, August 1992, *107* (3), 921–58.

Bergemann, Dirk and Juuso Välimäki, "Information Acquisition and Efficient Mechanism Design," *Econometrica*, May 2002, *70* (3), 1007–33.

Bernard, Andrew B. and J. Bradford Jensen, "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Economics*, February 1999, *47* (1), 1–25.

Black, Sandra E. and Lisa M. Lynch, "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity," *Review of Economics and Statistics*, August 2001, *83* (3), 434–45.

Bonelli, Regis, "A Note on Foreign Direct Investment and Industrial Competitiveness in Brazil," *Oxford Development Studies*, October 1999, *27* (3), 305–27.

____, Pedro da Motta Veiga, and Adriana Fernandes de Brito, "As Políticas Industrial e de Comércio Exterior no Brasil: Rumos e Indefinições," *IPEA Texto para Discussão*, November 1997, *527.* Instituto de Pesquisa Econômica Aplicada, Rio de Janeiro.

Boone, Jan, "Competitive Pressure: The Effects on Investments in Product and Process Innovation," *RAND Journal of Economics*, Autumn 2000, *31* (3), 549–69.

Brown, Lawrence D., *Fundamentals of Statistical Exponential Families, with Applications in Statistical Decision Theory*, Vol. 9 of *Lecture Notes–Monograph Series*, Hayward, California: Institute of Mathematical Statistics, 1986.

Browne, Sid, "The Return on Investment From Proportional Portfolio Strategies," *Columbia University PaineWebber Working Paper*, June 1995, *95/09.*

Burguet, Roberto and Xavier Vives, "Social Learning and Costly Information Acquisition," *Economic Theory*, January 2000, *15* (1), 185–205.

Calvo, Guillermo A. and Enrique G. Mendoza, "Rational Contagion and the Globalization of Securities Markets," *Journal of International Economics*, June 2000, *51* (1), 79–113.

Caplin, Andrew and John V. Leahy, "Business as Usual, Market Crashes, and Wisdom after the Fact," *American Economic Review*, June 1994, *84* (3), 548–65.

Cavalcanti Ferreira, Pedro and José Luis Rossi Júnior, "Evolução da Produtividade Industrial Brasileira e Abertura Comercial," *Pesquisa e Planejamento Econômico*, April 1999, *29* (1), 1–36.

Caves, Richard E., "Industrial Organization and New Findings on the Turnover and Mobility of Firms," *Journal of Economic Literature*, December 1998, *36* (4), 1947–82.

Citanna, Alessandro and Antonio Villanacci, "Existence and Regularity of Partially Revealing Rational Expectations Equilibrium in Finite Economies," *Journal of Mathematical Economics*, August 2000, *34* (1), 1–26.

___ **and** ___ , "Incomplete Markets, Allocative Efficiency, and the Information Revealed by Prices," *Journal of Economic Theory*, February 2000, *90* (2), 222–53.

Clarke, Richard N., "Collusion and the Incentives for Information Sharing," *Bell Journal of Economics*, Autumn 1983, *14* (2), 383–94.

Clerides, Sofronis K., Saul Lach, and James R. Tybout, "Is Learning by Exporting Important? Micro-dynamic Evidence from Colombia, Mexico, and Morocco," *Quarterly Journal of Economics*, August 1998, *113* (3), 903–47.

Cox, David and Richard Harris, "Trade Liberalization and Industrial Organization: Some Estimates for Canada," *Journal of Political Economy*, February 1985, *93* (1), 115–45.

da Motta Veiga, Pedro, "Export Promotion Policies in Brazil: Background, Current Situation and Main Trends," *Integration and Trade*, January-August 1998, *2* (4-5), 87–113.

Davis, Ronald E., "Certainty-Equivalent Models for Portfolio Optimization Using Exponential Utility and Gamma-Distributed Returns," in Kenneth D. Lawrence, John B. Jr. Guerard, and

Gary R. Reeves, eds., *Advances in mathematical programming and financial planning*, Vol. 3, Greenwich, CT and London: JAI Press, 1993, pp. 69–108.

DeGroot, Morris H., *Probability and Statistics*, 2nd ed., Reading, Massachusetts: Addison-Wesley, 1989.

Dixit, Avinash K. and Victor D. Norman, "Product Differentiation and Intraindustry Trade," in Avinash K. Dixit and Victor D. Norman, eds., *Theory of International Trade*, Cambridge, Cambridgeshire: Cambridge University Press, 1980.

Dubey, Pradeep, John Geanakoplos, and Martin Shubik, "The Revelation of Information in Strategic Market Games: A Critique of Rational Expectations Equilibrium," *Journal of Mathematical Economics*, 1987, *16* (2), 105–37.

Easley, David and Maureen O'Hara, "Time and the Process of Security Price Adjustment," *Journal of Finance*, June 1992, *47* (2), 576–605.

Eicher, Theo S., "Trade, Development and Converging Growth Rates: Dynamic Gains from Trade Reconsidered," *Journal of International Economics*, June 1999, *48* (1), 179–98.

Einy, Ezra, Diego Moreno, and Benyamin Shitovitz, "Rational Expectations Equilibria and the Ex-post Core of an Economy with Asymmetric Information," *Journal of Mathematical Economics*, December 2000, *34* (4), 527–35.

Ericson, Richard and Ariel Pakes, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, January 1995, *62* (1), 53–82.

Fama, Eugene F., "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, May 1970, *25* (2), 383–417.

Feenstra, Robert C., James R. Markusen, and William Zeile, "Accounting for Growth with New Inputs: Theory and Evidence," *American Economic Review*, May 1992, *82* (2), 415–21.

Fernandes, Ana M., "Trade Policy, Trade Volumes and Plant-level Productivity in Colombian Manufacturing Industries," November 2001. Yale University, Mimeograph.

Foster, F. Douglas and S. Viswanathan, "The Effect of Public Information and Competition on Trading Volume and Price Volatility," *Review of Financial Studies*, 1993, *6* (1), 23–56.

——— and ———, "Strategic Trading When Agents Forecast the Forecasts of Others," *Journal of Finance*, September 1996, *51* (4), 1437–78.

Funk, Mark and Jack Strauss, "The Long-Run Relationship between Productivity and Capital," *Economics Letters*, November 2000, *69* (2), 213–17.

Fuss, Melvyn, Daniel L. McFadden, and Yair Mundlak, "A Survey of Functional Forms in the Economic Analysis of Production," in Melvyn Fuss and Daniel L. McFadden, eds., *The Theory of Production*, Vol. 1 of *Production Economics: A Dual Approach to Theory and Applications*, Amsterdam: North-Holland, 1978, chapter II.4, pp. 219–68.

Gal-Or, Esther, "Information Sharing in Oligopoly," *Econometrica*, March 1985, *53* (2), 329–43.

Goh, Ai Ting and Jacques Olivier, "Learning by Doing, Trade in Capital Goods and Growth," *Journal of International Economics*, March 2002, *56* (2), 411–44.

Gomes, Victor, "Fatos Sobre a Produtividade," May 2001. IPEA-DF (Instituto de Pesquisa Econômica Aplicada), Brasília, Mimeograph.

Grossman, Sanford J., "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information," *Journal of Finance*, May 1976, *31* (2), 573–85.

——— **and Joseph E. Stiglitz**, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, June 1980, *70* (3), 393–408.

Harrison, Ann E., "Productivity, Imperfect Competition and Trade Reform: Theory and Evidence," *Journal of International Economics*, February 1994, *36* (1-2), 53–73.

——— **and Gordon Hanson**, "Who Gains from Trade Reform? Some Remaining Puzzles," *Journal of Development Economics*, June 1999, *59* (1), 125–54.

Hay, Donald A., "The Post-1990 Brazilian Trade Liberalisation and the Performance of Large Manufacturing Firms: Productivity, Market Share and Profits," *Economic Journal*, July 2001, *111* (473), 620–41.

Hellwig, Martin F., "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, June 1980, *22* (3), 477–98.

———, "Rational Expectations Equilibrium with Conditioning on Past Prices: A Mean-Variance Example," *Journal of Economic Theory*, April 1982, *26* (2), 279–312.

Hermalin, Benjamin E., "The Effects of Competition on Executive Behavior," *RAND Journal of Economics*, Autumn 1992, *23* (3), 350–65.

Holden, Craig W. and Avanidhar Subrahmanyam, "Risk Aversion, Liquidity, and Endogenous Short Horizons," *Review of Financial Studies*, Summer 1996, *9* (2), 691–722.

Hopenhayn, Hugo A., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, September 1992, *60* (5), 1127–50.

Horta, Maria Helena, Guida Piani, and Honório Kume, "A Política Cambial e Comercial," in IPEA, ed., *Perspectivas da Economia Brasileira 1992*, Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada, 1991, pp. 59–80.

IBGE, *Pesquisa Industrial Anual, Coleta Complementar: Manual de Treinamento.* Fundação Instituto Brasileiro de Geografia e Estatística (IBGE), Diretoria de Pesquisas (DPE), Departamento de Indústria (DEIND), Rio de Janeiro: Secretaria de Planejamento e Coordenação da Presidência da República, 1986.

——, *Pesquisa Industrial Anual, Coleta Especial: Manual do Agente.* Fundação Instituto Brasileiro de Geografia e Estatística (IBGE), Diretoria de Pesquisas (DPE), Departamento de Indústria (DEIND), Rio de Janeiro: Secretaria de Planejamento e Coordenação da Presidência da República, 1986.

——, *Estatísticas Históricas do Brasil. Séries Econômicas, Demográficas e Sociais, 1550-1988*, 2 ed., Vol. 3 of *Séries Estatísticas Retrospectivas*, Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística, 1990.

——, *Pesquisa Industrial Anual: Manual do Técnico de Pesquisa.* Fundação Instituto Brasileiro de Geografia e Estatística (IBGE), Diretoria de Pesquisas (DPE), Departamento de Indústria (DEIND), Rio de Janeiro: Secretaria de Planejamento e Coordenação da Presidência da República, 1994.

——, *Pesquisa Industrial Anual: Instruções para o Preenchimento do Questionário Completo de Empresa e Unidades Locais.* Fundação Instituto Brasileiro de Geografia e Estatística (IBGE), Diretoria de Pesquisas (DPE), Departamento de Indústria (DEIND), Rio de Janeiro: Secretaria de Planejamento e Coordenação da Presidência da República, 1996.

IOB, "Tabelas de Índices de Preços," in "Calendário Objetivo de Obrigações e Tabelas Práticas," Vol. Decembro de 2000, Brasília: IOB (Informações Objetivas), 2000, pp. 85–89.

Jackson, Matthew O., "Equilibrium, Price Formation, and the Value of Private Information," *Review of Financial Studies*, 1991, *4* (1), 1–16.

——— **and James Peck**, "Asymmetric Information in a Competitive Market Game: Reexamining the Implications of Rational Expectations," *Economic Theory*, May 1999, *13* (3), 603–28.

Jeffreys, Harold, *Theory of Probability.* The International series of monographs on physics, 1st ed., London: Clarendon Press, 1939.

Jordan, James S., "The Generic Existence of Rational Expectations Equilibrium in the Higher Dimensional Case," *Journal of Economic Theory*, April 1982, *26* (2), 224–43.

Keller, Wolfgang, "Do Trade Patterns and Technology Flows Affect Productivity Growth?," *World Bank Economic Review*, January 2000, *14* (1), 17–47.

Kim, Euysung, "Trade Liberalization and Productivity Growth in Korean Manufacturing Industries: Price Protection, Market Power, and Scale Efficiency," *Journal of Development Economics*, June 2000, *62* (1), 55–83.

Klette, Tor Jakob and Zvi Griliches, "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous," *Journal of Applied Econometrics*, July 1996, *11* (4), 343–61.

Knight, John L., Stephen E. Satchell, and Kien C. Tran, "Statistical Modelling of Asymmetric Risk in Asset Returns," *Birkbeck Discussion Paper*, June 1995, *05/95.* Birkbeck College, London.

Krugman, Paul R., "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 1980, *70* (4), 950–59.

——— **and Anthony J. Venables**, "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 1995, *110* (4), 857–80.

Kume, Honório, Guida Piani, and Carlos Frederico Bráz de Souza, "A Política Brasileira de Importação no Período 1987-98: Descrição e Avaliação," May 2000. IPEA (Instituto de Pesquisa Econômica Aplicada), Rio de Janeiro, Mimeograph.

Kyle, Albert S., "Continuous Auctions and Insider Trading," *Econometrica*, November 1985, *53* (6), 1315–35.

———, "Informed Speculation with Imperfect Competition," *Review of Economic Studies*, July 1989, *56* (3), 317–55.

Levinsohn, James, "Testing the Imports-as-Market-Discipline Hypothesis," *Journal of International Economics*, August 1993, *35* (1-2), 1–22.

—— **and Amil Petrin**, "Estimating Production Functions Using Inputs to Control for Unobservables," *NBER Working Paper*, August 2000, *W7819*.

Marschak, Jacob and William H. Jr. Andrews, "Random Simultaneous Equations and the Theory of Production," *Econometrica*, July-October 1944, *12* (3/4), 143–205.

Martin, Philippe and Gianmarco I. P. Ottaviano, "Growing Locations: Industry Location in a Model of Endogenous Growth," *European-Economic-Review*, 1999, *43* (2), 281–302.

Matsuyama, Kiminori, "Agricultural Productivity, Comparative Advantage, and Economic Growth," *Journal of Economic Theory*, December 1992, *58* (2), 317–34.

——, "Why Are There Rich and Poor Countries? Symmetry-Breaking in the World Economy," *Journal of the Japanese and International Economies*, December 1996, *10* (4), 419–39.

McFadden, Daniel L., "Conditional Logit Analysis of Qualitative Choice Behavior," in Paul Zarembka, ed., *Frontiers in econometrics*, Economic theory and mathematical economics, New York: Academic Press, 1974, pp. 105–142.

——, "Econometric Analysis of Qualitative Response Models," in Zvi Griliches and Michael D. Intriligator, eds., *Handbook of Econometrics*, Vol. 2, Amsterdam: Elsevier Science, 1984, pp. 1395–1457.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," November 1999. Harvard University, Mimeograph.

——, "Estimating Firm-level Productivity in Differentiated Product Industries," December 2000. Harvard University, Mimeograph.

Mesquita Moreira, Maurício and Paulo Guilherme Correa, "Abertura Comercial e Indústria: O Que se Pode Esperar e o Que Se Vem Obtendo," *Revista de Economia Política*, April-June 1997, *17* (2), 61–91.

Milgrom, Paul R., "Rational Expectations, Information Acquisition, and Competitive Bidding," *Econometrica*, June 1981, *49* (4), 921–43.

Moscarini, Giuseppe and Lones Smith, "The Optimal Level of Experimentation," *Econometrica*, November 2001, *69* (6), 1629–44.

Newey, Whitney K., "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, November 1994, *62* (6), 1349–1382.

Obstfeld, Maurice, "Models of Currency Crises with Self-Fulfilling Features," *European Economic Review*, 1996, *40* (3-5), 1037–47.

——— **and Kenneth Rogoff**, *Foundations of international macroeconomics*, Cambridge, Mass.: MIT Press, 1996.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, *64* (6), 1263–97.

Pavcnik, Nina, "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *NBER Working Paper*, August 2000, *W7852*.

Peletier, Ben D., "Terms of Trade Effects on Endogenous Growth Rates in LDCs," *Oxford Development Studies*, October 1998, *26* (3), 351–73.

Pietra, Tito and Paolo Siconolfi, "Fully Revealing Equilibria in Sequential Economies with Asset Markets," *Journal of Mathematical Economics*, March 1998, *29* (3), 211–23.

Popper, Helen A. and John D. Montgomery, "Information Sharing and Central Bank Intervention in the Foreign Exchange Market," *Journal of International Economics*, December 2001, *55* (2), 295–316.

Puga, Diego, "The Rise and Fall of Regional Inequalities," *European Economic Review*, 1999, *43* (2), 303–34.

Radner, Roy, "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices," *Econometrica*, May 1979, *47* (3), 655–78.

Rahi, Rohit, "Partially Revealing Rational Expectations Equilibria with Nominal Assets," *Journal of Mathematical Economics*, 1995, *24* (2), 137–46.

Raiffa, Howard and Robert Schlaifer, *Applied Statistical Decision Theory*. Studies in managerial economics, Boston, Massachusetts: Graduate School of Business Adminitration, Harvard University, 1961.

Raith, Michael, "A General Model of Information Sharing in Oligopoly," *Journal of Economic Theory*, October 1996, *71* (1), 260–88.

Ramos, Roberto Luís Olinto and Claudia Nessi Zonenschain, "The Performance of the Brazilian Imports and Exports Based on the System of National Accounts: 1980-1998," August 2000. IBGE Rio de Janeiro, Mimeograph.

—— **and** —— , "The Performance of the Brazilian Imports and Exports Based on the System of National Accounts: 1980-1998," August 2000. Paper presented at the 13th International Conference on Input Output Techniques. Macerata, Italy. August 21-25, 2000.

Rivera-Batiz, Luis A. and Paul M. Romer, "Economic Integration and Endogenous Growth," *Quarterly Journal of Economics*, May 1991, *106* (2), 531–55.

—— **and** —— , "International Trade with Endogenous Technological Change," *European Economic Review*, May 1991, *35* (4), 971–1001.

Robert, Christian P., *The Bayesian Choice: A Decision-Theoretic Motivation*, 2nd ed., New York: Springer, 1996.

Roberts, Mark J. and James R. Tybout, eds, *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*, Oxford: Oxford University Press for the World Bank, 1996.

Rodrigues, Agostinho Inácio, Edilton Pereira da Silva, and Sidney Ferro Barros, *A Nova Correção Monetária do Balanço: Lei n. 8200 de 28-6-91 Regulamentada pelo Decreto n. 332 de 4-11-91*, Brasília: IOB (Informações Objetivas), 1992.

Rodríguez, Francisco and Dani Rodrik, "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," in Ben Bernanke and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2000*, Vol. 15, Cambridge, Massachusetts: MIT Press, 2000.

Rodrik, Dani, *The New Global Economy and Developing Countries: Making Openness Work* Policy Essay No. 24, Washington, D.C.: Overseas Development Council, 1999.

Romer, David, "Rational Asset-Price Movements Without News," *American Economic Review*, December 1993, *83* (5), 1112–30.

—— , *Advanced macroeconomics.* McGraw-Hill advanced series in economics, New York: McGraw-Hill, 1996.

Romer, Paul M., "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 1986, *94* (5), 1002–37.

—— , "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–S102.

Routledge, Bryan R., "Adaptive Learning in Financial Markets," *Review of Financial Studies*, Winter 1999, *12* (5), 1165–1202.

Sachs, Jeffrey D. and Andrew Warner, "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity*, 1995, *1995* (1), 1–118.

Samuelson, Paul A., "The Pure Theory of Public Expenditure," *Review of Economics and Statistics*, November 1954, *36* (4), 387–89.

Schmidt, Klaus M., "Managerial Incentives and Product Market Competition," *Review of Economic Studies*, April 1997, *64* (2), 191–213.

Slaughter, Matthew J., "Per Capita Income Convergence and the Role of International Trade," *American Economic Review*, May 1997, *87* (2), 194–99.

Stokey, Nancy L., "The Volume and Composition of Trade Between Rich and Poor Countries," *Review of Economic Studies*, January 1991, *58* (1), 63–80.

Tybout, James R., "Plant- and Firm-level Evidence on "New" Trade Theories," *NBER Working Paper*, July 2001, *W8418.*

—— **and M. Daniel Westbrook**, "Trade Liberalization and the Dimensions of Efficiency Change in Mexican Manufacturing Industries," *Journal of International Economics*, August 1995, *39* (1-2), 53–78.

Verrecchia, Robert E., "Information Acquisition in a Noisy Rational Expectations Economy," *Econometrica*, November 1982, *50* (6), 1415–30.

Wang, Jiang, "A Model of Intertemporal Asset Prices under Asymmetric Information," *Review of Economic Studies*, April 1993, *60* (2), 249–82.

Weinhold, Diana and James Rauch, "Openness, Specialization, and Productivity Growth in Less Developed Countries," *Canadian Journal of Economics*, August 1999, *32* (4), 1009–1027.

Xie, Xin, "Contagion through Interactive Production and Dynamic Effects of Trade," *International Economic Review*, Feb 1999, *40* (1), 165–186.

Young, Alwyn, "Learning by Doing and the Dynamic Effects of International Trade," *Quarterly Journal of Economics*, 1991, *106* (2), 369–406.

Appendix A

Data appendix to chapter 3: *Pesquisa Industrial Anual* and complementary data sources

The present appendix is a detective's report. It documents the unglamorous but necessary efforts to construct the dataset of chapter 3. Brazil is one of few developing countries that surveys its industry systematically.[1]

The following section A.1 briefly describes the sampling method and the main types of variables in *Pesquisa Industrial Anual PIA*, the Brazil's annual industry survey. Section A.2 documents an analysis of the longitudinal relations between firms in this dataset. *PIA* traces in detail firm entries, exits, phases of suspended production (*mothballing*), mere changes of legal form, mergers, split-ups, spin-offs, and the like. However, this longitudinal aspect of *PIA* remains largely unexplored in economic research on Brazil to date. Section A.3 discusses ways to make the economic variables in *PIA* compatible over time and to correct for changes in surveying methods. In sections A.4 and A.5, I present methods to deflate the economic flow and stock variables—a task to be undertaken with much care since Brazil faces years of high inflation during the sampling period and changes legislation for the valuation of assets. Section A.6 discusses complementary data sources that are carry out the estimations of chapter 3.

At various instances, I mention short English variable names, sector definitions and regions in this appendix. Tables in section A.7 (p. 243; sectors), section A.8 (p. 248; regions) and sections A.9

[1] Among the developing countries with similar longitudinal industry databases are Chile, Colombia, Ivory Coast, South Korea, Mexico, Morocco, Taiwan, Turkey, and Venezuela (Levinsohn 1993, Roberts and Tybout, eds 1996, Clerides et al. 1998, Aw, Chung and Roberts 2000).

(p. 249; categories) and A.10 (p. 255; economic variables) list sectors, regions, variables and their descriptions. It is my hope that *PIA* be fruitful for microeconometric research in and on Brazil beyond this dissertation.

A.1 *PIA*—What It Contains, and What Not

Pesquisa Industrial Anual is an annual survey of Brazilian manufacturing firms and plants, conducted by the census bureau *IBGE* (*Fundação Instituto Brasileiro de Geografia e Estatística*). Even though the database's inception dates back to the seventies, a systematic and consistent sample of the Brazilian manufacturing sector is first assembled with the census of 1985. Until 1995, *PIA* surveys are based on the initial sample of 1986. Over the years, the surveys identify 10,507 legally established firms as potentially qualified for the *PIA* sample, out of which 9,151 firms exhibit manufacturing activity in at least one year. New firms enter the initial sample either because existing firms in the sample found them or because new firms are identified as sufficiently large 'greenfield' creations through a register at the labor ministry (*Relação Anual de Informações Sociais*). The firms in *PIA* between 1986 and 1995 are regarded representative of the medium-sized to large firms in their respective sectors. No survey exists for 1991 due to a federal austerity program that temporarily suspended the survey. The questionnaire is slightly reduced in 1992, but the sampling method continues unaltered. Today, the database between 1986 and 1995 is often referred to as *PIA velha* ("old *PIA*").

In 1996, the sampling method is changed to systematically include small and new-born firms. The complete *PIA nova* ("new *PIA*") includes roughly 40,000 firms. For the present purpose, however, I pay attention exclusively to those firms in *PIA nova* that are either present in *PIA velha*, too, or that are referenced as a longitudinally related firm by some firm in *PIA velha*. Exactly 5,278 firms satisfy this criterion.

A.1.1 Sample

PIA velha (1986-1990, 1992-1995) is a continuous sample of formally established, medium-sized to large Brazilian manufacturing firms for the years 1986 to 1990 and 1992 to 1995. The sample additionally embraces some medium-sized to large firms that are newly established between 1986 and 1993.

A firm is included in *PIA* only if at least half of its revenues stem from manufacturing and if it is formally registered as a tax payer with the Brazilian tax authorities (*Cadastro Geral do*

Table A.1: Layers of *PIA*

Year	Layer 1	Layer 2	Layer 3	Other[a]	*Subtotal*	Invalid[b]	*Total*
1986	789	5,975	0	21	6,785	950	7,735
1987	798	6,000	0	35	6,833	949	7,782
1988	802	6,190	1,347	70	8,409	1,229	9,638
1989	811	6,125	1,400	11	8,347	1,235	9,582
1990	808	6,077	1,322	42	8,249	1,238	9,487
1992	794	5,863	455	75	7,187	740	7,927
1993	791	5,438	467	109	6,805	565	7,370
1994	768	5,259	480	83	6,590	266	6,856
1995	761	5,076	437	83	6,357	192	6,549
Subtotal[c]					*9,151*	*1,356*	*10,507*
1996[d]	746	3,542	392	.	4,680	41	4,721
1997[d]	709	3,293	358	.	4,360	16	4,376
1998[d]	660	2,939	325	.	3,924	9	3,933
Subtotal[c]					*5,231*	*47*	*5,278*
Total[c]	*984*	*6,735*	*1,842*		*9,561*	*1,360*	*10,921*

[a]Firms entering due to the legal or economic change of a sample firm.

[b]Category of economic curriculum (`catlife`) is 9.3, 9.35, or 9.99. See appendix A.9.2.

[c]Number of firms that appear (appear and manufacture) in at least one year of *PIA*.

[d]The according layer is the firm's layer in 1995.

Contribuinte, *CGC*, at the time).[2] The sample of firms in *PIA velha* is constructed in 1986 from three layers:

- a non-random sample of the largest Brazilian manufacturers (called *coleta especial*),
- a random sample of medium-sized firms (*coleta complementar*), and
- a non-random selection of newly founded firms (*coleta de novos*).

A firm that ever enters *PIA velha* through one of the selection criteria remains in the *PIA velha* sample unless it is legally extinct. Moreover, if an existing firm in *PIA* reports the creation of a new firm as a subsidiary or spin-off, or the like, the according new firm is included in *PIA* too.

The criterion for inclusion in the first non-random layer is that the labor force of the firm either exceed an annual average of 1,000 employees in the census of 1985, or that its annual sales (*receita bruta*) in 1985 exceed a benchmark calculated in units of the governmentally imposed price index at the time (*OTN*). The cutoff value corresponds to roughly BRL 200 million in 1995 (around USD 200 million in 1995). Exactly 984 firms enter *PIA* through this layer. These firms make up for

[2]As a consequence of the 50-percent-manufacturer requirement, some manufacturing firms are disregarded. A large computer manufacturer in Brazil, for instance, engages in computer assembly, sales of services, and rental of equipment. It went unsampled in recent years because more than half of its sales stem from the latter two non-manufacturing activities.

about 9.6 per cent of all observations (firm-year combinations) between 1986 and 1995, and about 10.8 per cent of the 9,151 firms ever observed in operation in *PIA velha*.

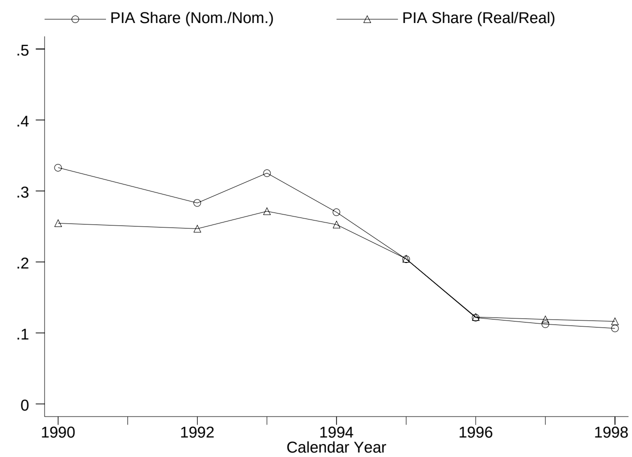
The second layer comprises randomly chosen firms that are identified during the census of 1985 and whose annual sales in 1985 exceed a cutoff value corresponding to roughly BRL 100,000 in 1995 (around USD 100,000 in 1995).

The third non-random layer of new-born firms comprises firms that emerge after the 1985 census. These firms are identified through the Brazilian labor ministry's register (*Relação Anual de Informações Sociais*). Only newly founded firms that surpass an annual average employment level of at least 100 persons are included. The inclusion process ends in 1993, however, so that greenfield creations are systematically observed only between 1986 and 1992. Even before 1993, the surveying method may not have been rigorously enforced at all times.

Due to the requirement that a firm be registered as a tax payer, firms in the so-called informal sector of the economy go unsampled by default. However, very few firms in the informal sector would attain a size that qualifies for one of the first two layers in *PIA velha*. So, every firm in *PIA* is uniquely identified by its tax number *CGC*.

PIA velha (1986-1990, 1992-1995) is complemented with those firms in *PIA nova* (1996-1998) that are longitudinally connected. This allows to trace about three quarters of the firms in *PIA velha* beyond 1995.

PIA nova is more representative of the Brazilian manufacturing sector as a whole then *PIA velha* was. There are only two layers in *PIA nova*. The first comprises a non-random sample of all medium-sized to large Brazilian manufacturers (more than 30 employees; about 27,500 firms). The second contains randomly selected small (at least 5 employees) to medium-sized manufacturers (12,500 firms). This may allow the construction of more systematic (unbalanced) firm panels in the future. For the purpose of constructing a continuous dataset beginning in 1986 and extending to the present, however, only a subsample of the firms in *PIA nova* seems adequate. *PIA velha* follows the principle that a firm once sampled be sampled again in every subsequent year unless extinct. In addition, greenfield creations do not make it into the *PIA* sample after 1993. This suggests a natural way to connect the two *PIA*s between 1995 and 1996. I select those firms in *PIA nova* that are either present in at least one year between 1986 and 1995, too, or that are longitudinally referenced by a firm in *PIA velha*. In *PIA nova* (until 1998) smaller firms are randomly sampled every year, and thus potentially randomly replaced every year. As a consequence, not all firms that are present in *PIA velha* reoccur in *PIA nova*. In fact, of the 6,549 firms present in *PIA velha* 1995, only 4,721 appear in *PIA nova* in 1996. In addition, initial problems in the register of firms for the *PIA nova* sample result in the omission of otherwise qualified firms.



Sources: Brazilian national accounts 1990-1998 (value added in manufacturing). Own calculations (total value added among manufacturers in *PIA*).

Figure A.1: Value added share of *PIA* in Brazilian manufacturing

This sample drop is a concern for estimation in chapter 3. However, the drop proves to be random and exogenous to the sample. Various treatments at different stages of the analysis in chapter 3—including the use of time indicators, period indicators and year indicators—do not show any significant impact on productivity estimates.

Table A.1 provides an overview of the size of the three layers in *PIA*. No economic information is available for the 'invalid' firms in the second-last column. However, their observations are kept in the sample to provide longitudinal information. These firms are initially identified as qualified, but, at the time when the *PIA* survey is conducted, they have gone out of business, turned out to be mainly non-manufacturing firms, or have been absorbed by another firm. As the exact results of the economic census of 1985 become known between 1986 and 1987, the sample of valid firms grows from around 6,800 firms in 1986 to about 8,400 in 1988. By 1992, it is down to roughly 7,300 firms again and drops to about 6,400 firms in 1995.

Figure A.1.1 shows the share of *PIA*'s firms in the Brazilian manufacturing sector as a whole (only longitudinally related firms are kept from *PIA nova*). *IBGE*'s national accounts office reports consistent value added figures for Brazilian manufacturing since 1990. The value added figures for *PIA* are constructed using the deflation methods discussed in section A.4. The medium-sized to large firms in *PIA* lose in market share relative to other Brazilian manufacturers. The decline occurs before the drop in sample size in 1996. In fact, the firms in *PIA* lose importance since 1993. Exit reduces the sample and becomes more frequent after trade liberalization in the early nineties. Also, Brazilian manufacturers that were smaller before are likely to gain in relative size.

A.1.2 Variables

Both *PIA velha* and *PIA nova* contain three main groups of variables: (a) Information about longitudinal relations across firms, (b) balance sheet information, and (c) economic information beyond the balance sheet. The according variables receive varying names and are kept in different ways over the years, but their individual content generally remains similar if not unaltered over time. Among the longitudinal information in group (a) are variables that indicate the state of activity of a firm in a given year (such as whether it operates all year, only part of the year, or exits) and its structural changes (such as whether it emerges from a pre-existing firm or whether it creates a spin-off firm itself, and the like). Variables in group (b) include cost, revenue, and profit information, detailed in a manner similar to a typical Brazilian income statement, and asset and liability figures until 1995, detailed in a manner similar to a typical Brazilian balance sheet. Variables in group (c) go beyond the balance sheet and income statement and include data such as investment flows, numbers of workers and employees, and a variable to indicate the origin of the firm's majority capital in *PIA velha*.

One of *PIA*s quite unique features is that it allows to distinguish between foreign and domestic machinery acquisitions for the years 1986 until 1995, and to distinguish between foreign and domestic intermediate goods purchases since 1996. In addition, two quite detailed variables indicating the state of a firm's economic activity allow to precisely trace the firms' operations over time so that researchers need not resort to assumptions about a firm's likely destiny when observations are missing. The variable indicating the origin of a firm's majority capital, however, is generally regarded as little informative. A main reason is that several firms in *PIA* are subsidiaries of Brazilian holdings which in turn are foreign-owned. Some of these subsidiary firms would interpret the variable in a strict sense and claim to be Brazilian-owned, while other firms would interpret the variable in a broader sense and indicate foreign ownership. As a consequence, the variable is imprecise.

The following section A.2 exploits the information of the longitudinal variables in group (a) in order to follow firms over time in an unbalanced panel and to calculate the economic age of firms. Sections A.3, A.4, and A.5 are dedicated to constructing consistent economic variables from the variables in groups (b) and (c) over time, and to their respective correction for inflation.

A.2 Longitudinal Relations between Firms

The longitudinal relations between firms in *PIA* have not systematically analyzed so far. Most researchers choose to work with *PIA* at various aggregate levels but not at the firm or plant level, partly because access to the confidential firm or plant data is restricted to researchers who

Table A.2: State of Activity

State of Activity	*PIA velha*	*PIA nova*
in operation	1	1
in installation phase	2	2
suspended production part of year	3	3
extinct	4	4
suspended production all year	5	5
extinct in earlier year	6	6, 7
mainly non-manufacturing	7	8
other	8	9, 10, 11, 12, 13, 14

temporarily affiliate themselves with *IBGE* Rio de Janeiro, while *IBGE* shares data on aggregate levels. Second, throughout the period from 1986 to 1998, the internal data analysis and critique at *IBGE*'s industry division is designed to check data for their consistency within a given year but not to check their consistency across years.

The present section documents the use of longitudinal information in *PIA velha* and *PIA nova*. This information serves as a key component for the construction of an unbalanced firm panel. It needs to be known if and when a firm exits, whether data for future years are simply missing or whether a firm chooses to temporarily suspend production, whether an exiting firm survives in different legal form or really stops producing, and the like. As a side product of this information, the economic age of a firm is inferred.

A.2.1 States of activity and types of change

Both *PIA velha* and *PIA nova* contain two variables that are intended to reveal precise information about the economic and legal state of a firm. The *state of activity* (`state` *situação cadastral* indicates whether a firm operates in a given year. Table A.2 summarizes the level of detail of the according variable in *PIA velha* and *PIA nova*.

A second variable can be translated as *structural change* or *change of economic/legal status* (*mudanças estruturais*; variable: `change`). It records changes to the firm's legal and economic status. The classification is considerably simplified in *PIA nova* (variable: `'change'`) as compared to the earlier *PIA*. In order to make this variable compatible between *PIA velha* and *PIA nova*, algorithms as indicated in table A.3 are applied. Two further variables indicate the month and year at which the change occurs (`chmon` and `chyr`). Knowledge about this timing is important to properly deflate the according economic variables.

PIA velha also documents one so-called *tax number link* (*CGC de ligação*) and *PIA nova* up to three such *tax number links*. They serve to connect firms over time so that successors and

Table A.3: Change of Legal or Economic Status

Change of Legal/Economic Status	*PIA velha* `change`	*PIA nova* '`change`'	`state`	additional
no change	.	.		
merger	1	1		> 1 predec.
absorbed into other firm	2	3	4, 5	
absorbing other firm	3	3	1, 2, 3	
complete split-up into successor(s)	4	1		1 predec.
partial spin-off into existing firm	5	2		succes. old
partial spin-off into new firm	6	2		succes. born
dissolved	7	-	4, 6, 7	
(parts) rented out to other firm	8	(4)		
renting (parts of) other firm	9	(5)		
other[a]	10	6		

[a] As explained in the according manuals, the category 'other' is systematically used in *PIA velha* (IBGE 1986a, IBGE 1986b). For example, it is generally assigned to the firm that arises from a merger. With *PIA nova*, the use of 'other' becomes restricted to otherwise unclassified cases (IBGE 1996).

predecessors are identified. A peculiar feature of these *tax number links* both in *PIA velha* and *PIA nova* is that they are used as connectors for the referencing firms as well as connectors for the firm that is being referenced. Hence, they change their meaning depending on the value that the variable `change` takes. Firms are asked to provide the year of their foundation in *PIA velha*, and the same information is available for firms in *PIA nova* through tax registers and *IBGE*'s own register of known Brazilian manufacturers (*Cadastro Básico de Seleção* for *PIA nova*). Column five ('additional') in table A.3 exploits these two types of additional information to make the variable *change of legal/economic status* (`change`) compatible between *PIA velha* and *PIA nova*.

A.2.2 Reclassifications and error corrections

Table A.4 summarizes my reclassifications and corrections for the variables *state of activity* (`state`) and *change of legal/economic status* (`change`). In the case of several firms, the variables `state` and `change` exhibit contradictory patterns over time. These conflicts are often hard to resolve. Therefore, I generally choose to sort firms out whose longitudinal data exhibit such contradictions. However, some unnecessarily vague classifications or obvious mistakes are corrected. Whereas the upper four reclassifications seem to be due corrections, the reclassifications in the lower part of table A.4 seem justified but not necessary. A telling example may be the reclassification to `state`:=5 in line 7. The combination of economic circumstances (`state`=8, `change`=8, no sales) suggests that a firm rents its equipment to another firm while it realizes no sales of its own—hence, it suspended own production in fact. (There are 44 such observations in *PIA*.)

Table A.4: Reclassifications and Error Corrections

Correction	if, in a given year,
`change`:=2	`change`=3, and `state`=4 or 6
`change`:=3	`change`=2, `state`=1, and firm continuously present in *PIA*
`state`:=2	`state`=5 or 8, no sales, and year=first year of appearance
`state`:=5	`state`=8, no successor, and `state`=3, 4, 5, or 6 in following year
`state`:=4	`state`=8, positive sales, and `change`=1, 2, 4, or 7
`state`:=4	`state`=1, `change`=10, and year is last year
`state`:=5	`state`=8, `change`=8, and no sales
`state`:=5	`state`=8, `change` empty, no successor, and no sales
`state`:=6	`state`=5 and year=last year of appearance (and before 1998)
`state`:=6	`state`=8, no sales, firm has successor, and `change`=1, 2, 4, or 7
`state`:=6	`state`=8, no sales, no successor, and `change` empty

More often than necessary, firms choose the category 'other' as `state` or `change` to classify the type of change they undergo. Natural reclassifications are listed in table A.4. Finally, firms may sometimes have alleged incentives to misrepresent the category of `change`. An example is that a firm merely alters its legal form with no economic consequences for the production process—in order to realize advantages in taxation or at financial markets, say. Especially when taxation is concerned, this firm would typically claim in the questionnaire that its predecessor is extinct (`change`=7) without any remainders, but would still provide the *tax number link* to this predecessor and possibly add hand-written observations (while `state`=8). The correct category of `change` would be 'dismantled into successor' (`change`=4). The reclassifications in table A.4 take this and similar misrepresentations into account.[3]

The information in `chmon` and `chyr` indicates in what month of a year the recorded `change` occurs. Firms often report a month of change `chmon` in later years that is different from the `chmon` in earlier years, or they do not report a month of change initially but provide one later. Errors in this variable may slightly affect the method of inflation correction proposed for flow variables in section A.4.1. In general, I correct the information in `chmon` so that the longest justifiable survival time of the firm results, that is to use the latest exit month reported when information is contradictory. This procedure makes errors from the correction method in section A.4.1 the least likely. Also, omissions or errors in the variable `chyr` affect the construction of the 'family tree' of firms, that is the 'parent-child' relations between firms (see section A.2.4). I insert missing information in `chyr` if this information is consistently provided in later years.

[3]The alternative of manually reviewing several hundreds to thousands of hand-written observations for every year in *PIA* seems an undue effort.

Table A.5: Proper Parents and Children for *PIA*'s Family Tree

Properly referencing firms (proper parent), through *tax number link(s)*
state=4 or 6, change≠10, and year=effective exit year
state=4 or 6, and change=1, 2, 4, or 7
state=1, change=10, year=effective exit year, and `indcor` records change[a]
Properly referenced firms (proper child), through *tax number link(s)*
change=3, and successor firm identified
change=10,[b] and successor firm identified

[a]The variable `indcor` is an indicator variable only available for *PIA velha* between 1992 and 1995. It states whether a firm merely changes its tax number.

[b]See footnote *a* in table A.3.

A.2.3 Effective suspension and exit times

It proves helpful to know the effective times of a firm's exit and the exact beginning of periods of temporarily suspend production (mothballing). In *PIA*, recorded exit years are often preceded by missing years or years with special observations but no sales (`state`=8 or `change`=10, 'other'). Similarly, years of suspended production are often surrounded with periods of missing years or years with special observations. I calculate the effective exit (suspension) year for every firm (`effextyr` and `fstsusyr`, respectively) as the earliest year preceding an observed exit (suspension) year after which no proper year is observed until exit (suspension) is recorded indeed.

A.2.4 Identifying longitudinal links in a 'family tree'

A simple way to trace firms over time is to construct a family tree that records the parent-child connections between firms. This approach is briefly outlined here. The following subsection A.2.5 below is devoted to the more involved task of classifying the economic curriculum of firms, whether they are connected to other firms or not.

A family tree of firms and their predecessors can be arranged in a list where the lines are of the form:

line 1	Firm D	←	Firm B	←	Firm A
line 2	Firm D	←	Firm C		
...					

In this setup, the oldest forefather of a firm lies the farthest to the east, and all children are listed below each other in the west.

The particular way in which *PIA*'s longitudinal information is arranged suggests to build this family tree up from two sides. The *tax number link(s)* are used both with the referencing firm

Table A.6: Simplified Categories of Entry

Type of Entry (`catentsi`)	Original Category (App. A.9.1)	No. of Firms
1: old firm	all other categories	10,018
2: baby firm	2.1, 2.4, 4.11, 4.14, 8.4, 9.2	895
9: problem firm	9.1	8
Total		*10,921*

(the parent) and with the firm that is being referenced (the child). These *tax number link(s)* change their meaning according to the value that the variable `change` takes. The upper part of table A.5 shows which firms are selected as properly referencing firms (proper parents), thus building-up the family tree from the east. Similarly, table A.5's lower part shows firms that are selected as being properly referenced (proper children) so that the table is built-up simultaneously from the west.

Due to the arrangement of the longitudinal information in *PIA* (especially in *PIA velha*), several but by far not all entries occur twice in *PIA*—justifying the double build-up effort. As it turns out, double entries in *PIA*'s family tree never contain conflicting information for *PIA*—a reassuring fact given that *PIA*'s data between 1986 and 1998 are generally not submitted to dynamic checks across years. The maximal number of 'generations' in the family tree for *PIA* (1986-1998) are three parent-child relations. However, some alleged predecessors are not contained in *PIA* (indicated by the variables `pr1noshw`, `pr2noshw`, and `pr3noshw`). Overall, about a tenth of all firms in *PIA* (1,099 firms) are identified as 'children' of some parent in or out of sample.

A.2.5 A firm's 'economic curriculum'

For the construction of an unbalanced panel and its econometric application, we need to know both why a firm enters the data set and why it drops out. A firm that enters the data as a mere legal successor of a previously surveyed firm has to be treated differently from a greenfield creation. Similarly, a firm that is dismantled but lives on in many successor firms is to be clearly distinguished

Table A.7: Simplified Categories of Exit and Suspended Production

Type of Life (`catlifsi`)	Original Category (App. A.9.2)	No. of Firms
1: always healthy	all other categories	6,188
2: suspends, returns, never exits	3, 3.1, 5.3, 5.311-5.313, 8	727
3: suspends, returns, exits later	3.2, 5.314, 5.32	410
4: exits	1.4, 2, 5.14, 5.2	2,115
6: reclassification possible	9.1, 9.15, 9.2, 9.35	125
9: problem firm	9.3, 9.99	1,356
Total		*10,921*

Table A.8: Simplified Entry and Life Categories, combined

Type of Life (`catlifsi`)	Type of Entry (`catentsi`) 1: old firm	2: baby	9: problem	*Total*
1: always healthy	5,544	642	2	*6,188*
2: suspends, returns	657	70	0	*727*
3: suspends, exits later	390	20	0	*410*
4: exits	1,969	144	2	*2,115*
6: to be reclassified	111	6	4	*121*
9: problem	0	0	0	*0*
Total	*8,671*	*882*	*8*	*9,561*

from a firm that stops producing for good. The variables `state` and `change`, together with complementary information, allow for a fairly precise characterization of the economic curriculum of a firm in *PIA*.

For many econometric applications a few categories of entry ('old' or 'baby', say) and exit ('healthy', 'suspended', or 'shut-down', say) may suffice. However, to arrive at such simple categories, firms generally need to be sorted according to a more detailed roster first. Appendices A.9.1 (p. 250) and A.9.2 (p. 252) document the two fine rosters for entry and exit that are used.

Tables A.6 and A.7 present a condensed classification, derived from the detailed categories in appendices A.9.1 and A.9.2, respectively. Some classifications are certainly debatable. For example, a spin-off firm created by an existing firm is considered a baby firm with zero economic age in the categorization of table A.6. However, it might also be justifiable to categorize a spin-off as old firm with the age of the economic predecessor. Table A.6 treats firms that emerge from a complete split-up of their predecessor in this latter way, for example. The idea is that spin-offs are founded to stand alone and gain experience on their own, moving away quickly from the parent firm's original knowledge, whereas successor firms from a complete split-up may benefit more from the initial knowledge incorporated in their plants, continuing the business of their predecessor. Clearly, this classification is a judgement call.[4]

Table A.8 summarizes tables A.6 and A.7 and shows the number of firms in *PIA* that are observed with positive manufacturing sales in at least one year.

Table A.9: Effective Creation Time

Algorithm to find effective **legal founding year** (`effborn`)
Set `effborn` to registration year in *IBGE*'s most recent register (*Cad. Básico*)
Replace `effborn` by reported year in *PIA velha* **if** `effborn` is 1965 or 1966
Replace `effborn` by year of first appearance in *PIA* **if** first appearance earlier
(Sort firm out **if** registration year later than first appearance in *PIA*)

Algorithm to find effective **economic founding year** (`econborn`)
Set `econborn` to `effborn`
Replace `econborn` by founding year of predecessor **if** firm emerges from split-up or spin-off
Replace `econborn` by founding year of absorbed predecessor **if** predec. large
(avg. labor force of predec. at least two thirds of firm's avg. labor force)

A.2.6 A firm's economic age

The economic age of a firm is of interest for its own sake and can help check longitudinal relations in addition. There are several sources to infer the age of a firm in *PIA*. The *PIA velha* questionnaire asks for the firm's founding year; tax registers and *IBGE*'s own register of known manufacturers (*Cadastro Básico de Seleção*) record the year of a firm's legal creation; and the year of first appearance in *PIA* together with an observed `state` 'in installation' may be indicative. As it turns out, these sources contain partly contradictory information. Common reasons are that firms only register their creation at the tax roll with a delay and that some firms only enter an approximate founding year in the questionnaire. In addition, recent copies of the tax register contain truncated information in the year 1966; that is, firms founded before 1966 are recorded as created in 1966. There are several possible sets of criteria to infer a firm's founding year from this conflicting information. The set of criteria applied here is presented in the upper half of table A.9.

The founding year may not reflect the true economic age of a firm. For example, reasons of taxation or legal causes may induce a firm to change its legal status while it remains the same economic entity. Clearly, such a firm should be considered older than the registration year of the most recent tax number. The founding year corresponds to the 'legal age' of a firm, whereas its 'economic age' is determined by the impact of its predecessors. Again, there are several criteria to infer an adequate economic age of firms in *PIA*. In one way or another, they all make use of the information in a firm's family tree as discussed in subsection A.2.4 above. The lower part in table A.9 describes a possible algorithm.

[4] These difficulties in classification do not only arise in the case of firms. One encounters similar problems with plants. If an existing firm opens a new plant, for example, it is not clear whether the new plant should be assigned the age of the founding firm (as it receives human capital and implicit knowledge transfers) or be counted with zero age (as it starts a new production process).

A.2.7 Regional classifications

The variables `region` and `uf` indicate the location of the legal headquarters of a firm (see appendix A.8 for an overview of Brazil's regions). The location of the headquarters need not coincide with the region of a firm's main economic activity or value creation. In principle, a value-added based reclassification of the variables `region` and `uf` could be inferred from plant-level information in *PIA* for a number of firms, but not for all firms since there is no complete overlap between plant-level and firm data. The regional variables exhibit strange observations in a few instances. Entries below '1' or above '5' in the variable `region` are set to missing. In some cases, missing values for the variable region are inferred from `uf`, the more detailed variable. Finally, I classify the region of a firm to the one in the preceding or following year if an observation is missing, depending on whether a change of region occurs or not.

A.2.8 Sector classifications

Firms in *PIA velha* are classified into sectors according to *Nível 100* (for a description of sectors see appendix A.7). In *PIA nova* the sector classification is changed to *CNAE* (*Classificação Nacional de Atividades Empresariais*). Since *CNAE* is more detailed, firms in *PIA nova* are re-classified to *Nível 100* (see appendix A.7 for a translation key). However, there is a break between *PIA velha* and *PIA nova*. Many firms apparently change sectors between 1995 and 1996. This may have to do with the fact that outdated firm classifications in *PIA velha* are corrected in *PIA nova*. As a consequence, adjustments over time may be in place. If one wants to use the years 1992 through 1998, for instance, it may be worthwhile to only use sector classifications from *PIA nova*. However, since I choose to cover the entire period from 1986 through 1998, sector classifications of *PIA velha* seem to be more adequate and are used. A downside of these adjustments is, of course, that changes in a firm's product range are disregarded.

A.3 Compatibility of Economic Variables Over Time

Between 1986 and 1998, *PIA* suffers two structural breaks. The questionnaire is slightly simplified and partly downsized in 1992. In 1996, with the creation of *PIA nova*, several economic variables drop out, some few are added, and the aggregation of variables from the balance sheet and income statement changes. To obtain time consistent economic variables for the entire period from 1986 to 1998, a few adjustments are in place.

A.3.1 Time-consistent economic variables

Table A.14 in appendix A.10 (p. 256) documents the manner in which I construct consistent economic variables. In the present section, I discuss main concerns.

Some changes in variable definitions are noteworthy. Gross sales, including taxes and subsidies, incorporate changes that are not due to market forces. Net sales are used instead. However, sales figures in *PIA velha* and *PIA nova* seem to be most compatible when gross sales are considered both between 1986 and 1995 (including export subsidies, credit subsidies such as *IPI*) and between 1996 and 1998 (including the usually small additional revenues from services). The according variable is named `grssales` in table A.14 (p. 256). At any level net of subsidies or service revenues, sales figures are not immediately compatible across *PIA velha* and *PIA nova*—due to a re-grouping of the variable definitions in the questionnaire in *PIA nova*. However, there is an alternative. Make the assumption that export and credit subsidies as well as service revenues move in fixed proportion to total sales within any given year. Taxes are generally calculated in fixed proportions of total sales. Then I can calculate adjusted net sales as the fraction of net sales that is due to other economic activity than taxes, subsidies and service revenues. The according variable in table A.14 (p. 256) is `sales`, which is used as one component of the production proxy in chapter 3.

The redefined salary variable in the *PIA nova* questionnaire makes a similar effort necessary for wages. I distribute the (extra position) of 'gratuities and bonuses' linearly between blue and white-collar salaries for *PIA velha*. These gratuities and bonuses are included in the respective salary variables in *PIA nova*.

In *PIA nova*, computer acquisitions are lumped together with other acquisitions. The variable `acqother` reflects the correct sum for all years 1986-1998 while `acqcomp` gives the value of computer acquisitions between 1986 and 1995. The same classification applies to the asset retirements of computers and other capital goods (`aslother` and `aslcomp`).

For reasons hard to understand today, intermediate goods acquisitions did not receive a position of their own in the *PIA velha* questionnaire. The best proxy for intermediate goods acquisitions is the variable called 'other costs and expenditures' (*outros custos e despesas*). This weakness of *PIA velha* makes it necessary to construct a similar (and equally noisy) variable for *PIA nova*. I add purchases of intermediate goods'(*compras de matérias-primas, materiais auxiliares e componentes*), the total of combustibles, electric energy, and services consumption (*consumo de combustíveis*, *compra de energia elétrica*, *consumo de peças*, *serviços industriais*, and *serviços de manutenção*) as well as shipping costs (*fretes e carretos*) and other operational cost (*demais custos e despesas operacionais*) in *PIA nova*.

The variables `wagetop` and `wagewh`—representing the salaries of top managers (firm owners) and white-collar employees, respectively—cannot be made exactly compatible between *PIA velha* and *PIA nova*. The reason is that *PIA velha* and *PIA nova* treat upper-level managers (*diretores*) in a different manner. While *PIA velha* includes upper-level managers' salaries in the variable `wagetop` (together with top managers and firm owners), *PIA nova* includes these managers' salaries in `wagewh` (together with employees).

During the first years of *PIA velha* (1986-1990), firms are asked to present the steps of their asset revaluation under inflation in the *PIA* questionnaire (emphcorreção monetária). However, the according fields in the questionnaire are arranged in a contradictory manner (asset acquisitions, for example, appear before the monetary correction column, rendering it unclear at what stage the appropriate correction should be presented). This and further problems made the variables never pass the data critique. Consequently, the fields are dropped after 1992. I include only stock variables such as `aspmasum` from these fields in the dataset. Similarly, variables such as final stocks of vehicles and computers could be included. Since they reflect final values after all monetary corrections, these variables are not likely to suffer from contradictory monetary correction steps.

A time-consistent variable `profit` is constructed for profits *before* taxes. No consistent series of profits *after* taxes can be derived because questionnaires in *PIA velha* (1986-95) and *PIA nova* (from 1996 on) differ. Before 1996, the reported profit figure is profit *after* tax and workers' participation, and the latter two costs are reported. Since 1996, the reported profit figure is profit *before* tax and workers' participation, while the latter two costs are not reported. So, only a series of profits *before* taxes is constructed that is consistent in this respect. For this purpose, one can add back anticipated taxes and workers' participation to after-tax profits in *PIA velha*. In a strict sense, the proposed profit series still suffers from a slight incompatibility for the years 1989 and 1990. The reason is a legal change in 1988 that is only accounted for in the *PIA* questionnaire after 1990. Social contributions under *lei 7689 de 15/12/1988* reduce profits in addition to the tax payments from 1989 on. Only the questionnaires after 1991 include these payments explicitly. However, the reported costs of social contributions under *lei 7689 de 15/12/1988* are small on average (2.6 percent between 1992 and 1995) so that the implied error in the profit figures in 1989 and 1990 should be small. In addition, not given any other choice, firms in 1989 and 1990 are likely to report this cost under taxes so that it would be accounted for.

Finally, observe that both the variable `difstock` and the variable `intmdif` are calculated departing from cost information in the income statement. Therefore, a positive value means a *decrease* in stocks. These variables are used to arrive at the full production on the output side and the full use of intermediates and materials at the input side.

Table A.10: Rebasing to Brazilian Real as Common Currency

Year[a]	Currency[b]	in BRL (July 1994)[c]	change in[d]
(1985)	*Cruzeiros*	1/(2.75*1,000,000,000,000)	
1986	*Mil Cruzados*	1/(2.75*1,000,000)	March 1, 1986
1987	*Mil Cruzados*	1/(2.75*1,000,000)	
1988	*Mil Cruzados*	1/(2.75*1,000,000)	
1989	*Mil Cruzados Novos*	1/(2.75*1,000)	
1990	*Mil Cruzeiros*	1/(2.75*1,000)	
(1991)	*Mil Cruzeiros*	1/(2.75*1,000)	
1992	*Mil Cruzeiros*	1/(2.75*1,000)	
1993	*Mil Cruzeiros Reais*	1/(2.75*1,000)	August 1, 1993
1994	*Reais* (BRL)	1	July 1, 1994
1995	*Reais* (BRL)	1	
1996	*Reais* (BRL)	1	
1997	*Reais* (BRL)	1	
1998	*Reais* (BRL)	1	

[a] December of the year. *PIA* is based on end of year values.

[b] As used in the *PIA* micro-data base. *Mil* means 1,000.

[c] The factors need not apply to published aggregate figures from *PIA*.

[d] Applicable to monthly deflators.

A.3.2 Missing values in *PIA velha*

In *PIA velha*, zero values of observations cannot be distinguished from missing values. Depending on the type of variable, I choose different procedures to decide which value should be regarded as missing and which one as zero. In the case of sales, for instance, it is likely to make little difference whether a value is missing or zero. The firm is regarded as not in operation. However, when observations of gross investment are missing, as another example, it does matter whether a value is zero or missing. It also becomes harder to decide whether no investment is undertaken indeed or whether investment is incorrectly reported in the questionnaire. In this particular case, I consider a value of gross investment as zero when the according asset retirements figure is not missing, and as missing otherwise. Similar criteria are applied to other variables. *PIA nova* properly distinguishes between missing and zero values.

A.3.3 Rebasing to a common currency

During the sampling period of *PIA*, the Brazilian currency changes four times (but only twice the currency units are altered). All variables in the *PIA* database are in current currency of the according year. Table A.10 shows how the figures in *PIA* are rebased to one common year. The factors in table A.10 refer to the latest Brazilian currency *Real* (BRL, introduced in July 1994).

A.3.4 A comment on plant data in *PIA velha*

At the plant level (*unidade local*), several further precise variables are available in *PIA velha*: For example, the consumption of combustibles and electric energy in production and more precise information about the use of intermediate products. While it seems hard to break firm-level data (such as investment flows or the capital stock which are not directly observed at the plant level) down to the plant level, it might seem a natural extension of the dataset to aggregate the plant data into firm data and then use the more complete dataset. However, this approach proves little rewarding.

The sample of plants in *PIA velha* is constructed in a manner very similar to the sample of firms. The non-random part comprises the plants of the leading firms (in layer 1; see table A.1, p. 199). The random part, however, consists of plants that are randomly drawn themselves—independently of the firms that enter *PIA velha*. Therefore, only very few plants and firms overlap. As a consequence, a joint dataset of plant-level data, aggregated into firms, and merged firm-level data results in a sample of considerably less than 1,000 firms. Depending on how one counts firms with missing data, the usable sample may only comprise some 200 to 400 firms. In addition, these firms are concentrated in very few sectors. Compared to a sample of more than 9,500 firms, the little gain in additional information from merging plant-level and firm-level information does not seem justified.

A.4 Deflating Flow Variables

Brazil faces periods of extremely high inflation until the *Plano Real* finally succeeds in bringing down inflation in July 1994. The average annual inflation rate between January 1986 and December 1994 is 820 per cent (according to *INPC*), while the *Plano Real* brings inflation down to a yearly average of 8.8 per cent between January 1995 and December 1998 (*INPC*). As a result, the data, especially in *PIA velha*, need to be carefully corrected for inflation.

Firms in both *PIA velha* and *PIA nova* are asked to provide economic variables in the same manner as they would present the figures in their balance sheet or income statement. However, civil law and the according accounting orders of the federal government are often designed as if inflation did not exist. Moreover, several officially imposed price indices deliberately understate true inflation. Together, these two factors create substantial difficulties for the researcher to arrive at realistic real values of the variables. The legal stipulations affect flow and stock variables in quite different ways. I will discuss both groups of variables separately below.

In *PIA velha* and *PIA nova* firms are asked to report economic numbers referring to the calendar year of the survey. Firms whose business year does not coincide with the calendar year are required to adjust the numbers accordingly. The monetary correction for inflation has to be conducted following *Legislação Societária* (e.g. IBGE 1994). *PIA*'s instructions mandate explicitly that firms not apply *Correção Monetária Integral* ('complete monetary correction') which contain a set of rules for monetary adjustment of both flow and stock variables. Instead, firms are asked to follow *Legislação Societária* (see e.g. IBGE 1994, p. 48). Brazil's *Legislação Societária* is grounded in *Lei n. 6404 de 15-12-76.* This law and the according governmental orders, still in force as of 2001, prohibit the monetary correction of flow variables. The law does, however, specify procedures for revaluing assets under inflation.

A.4.1 Correcting for ignored inflation

Since Brazil's *Legislação Societária* does not allow to deflate flow variables, all economic variables in *PIA* that stem either from the firm's income statement or relate to salaries are simple sums of the firm's monthly (or possibly daily) figures. Under high inflation, a simple sum depresses the January values considerably and correctly represents just a about the (late) December values. There seems to be no direct way to recapture more precise inflation-adjusted figures. I therefore use the following approximation to a more realistic value for the flow variables.

Call the observed value of the respective flow variable in year t $\tilde{X}_t$. $\tilde{X}_t$ is the value reported by *PIA* but it reflects the wrong sum of not corrected nominal flow values. Similarly, call the correct real value of the firm's annual figure X_t. Suppose that the firm has a proper monthly accounting system and that it simply sums its monthly figures up to the annual figure, for which the *PIA* questionnaire asks. Suppose also that the monthly accounting system correctly adjusts for inflation over the course of the month. If one finally supposes that the firm's annual figures suffer from no seasonal fluctuations over the course of the year, the wrong annual value is

$$\tilde{X}_t = \frac{X_t}{12}\frac{\pi_{jan,t}}{\pi_{dec,t}} + \frac{X_t}{12}\frac{\pi_{feb,t}}{\pi_{dec,t}} + \ldots + \frac{X_t}{12}\frac{\pi_{dec,t}}{\pi_{dec,t}}, \tag{A.1}$$

where $\pi_{month,t}$ denotes the according monthly price index.

This equation says: If the annual figures are evenly distributed across months ($\frac{X_t}{12}$ is the same every month) then we can commit the same error as the firm had to commit when applying *Legislação Societária*. We can simply downsize the January figure by the inflation rate between January and December, downsize the February figure by the inflation rate between February and December, and so forth, and then sum all these inappropriate monthly figures up to the wrong annual figure $\tilde{X}_t$. This is the error that all firms in *PIA* are forced to commit when presenting their figures for flow variables. Of course, one can undo this error by solving (A.1) out for X_t. This yields

$$X_t = \frac{12 \cdot \pi_{dec,t}}{\pi_{jan,t} + \pi_{feb,t} + \ldots + \pi_{dec,t}} \tilde{X}_t. \tag{A.2}$$

Since we know or can construct appropriate price indices for all kinds of flow variables in *PIA*, we can apply equation (A.2) to every single flow variable in *PIA* and arrive at corrected annual values. These values come closer to a realistic real annual value than the raw number in *PIA* does—even if we have no reason to believe that there are no seasonal fluctuations over the year. I apply the correction of equation (A.2) to every flow variable in *PIA*. For firms that got out of business during a year, the variable `chmon` indicates the month of effective exit. I use the formula only up to the respective exit month. A remaining task is to find or construct appropriate monthly price indices for each flow variable.

A.4.2 Price indices for 1986-1998

Depending on the circumstance, either sector-specific or industry-wide price indices are more appropriate to deflate flow variables. In principle, the use of an industry-wide (or even economy-wide) price index has the benefit of maintaining the relative price structure across sectors, regions and time, while it supposedly captures all monetary effects on the price level. Moreover, industry-wide price indices avoid washing out relative price changes that stem from sector-specific quality improvements. However, in the case of Brazil mainly two practical concerns tend to wipe out the benefits of industry-wide price indices. At times of high inflation, the Brazilian federal government imposes price controls in various sectors that are more easily controlled or more prominent in consumers' minds, while it leaves other (usually the less concentrated) sectors unrestricted. As a consequence, not all prices keep pace with the growth of money supply and price distortions across sectors arise. Similarly, regional and sector-specific conditions (such as contract types, the concentration of industry, and the like) make the price adjustment to inflation less flexible in some sectors or regions, while it is more rapid and adequate elsewhere. These rather monetary factors are likely to distort price differences more strongly than real factors (such as quality, demand, or supply changes, say). As a consequence, sector-specific price indices seem more appropriate than industry-wide indices.

As a general conclusion, a sensitivity analysis with respect to differently deflated data seems key whenever working with *PIA* before 1994. Only a sensitivity analysis is likely to provide an adequate robustness check for the reliability of statistics and estimates, and an assessment of likely distortions through high inflation. Accordingly, the productivity estimation in chapter 3 is carried out for three alternative deflation methods.

Useful industry-wide price indices are *IPA-OG* (*Índice de Preços por Atacado–Oferta Global*, wholesale price index covering the entire economy including imports; by *FGV*), *IPA-INDTOT* (*Índice de Preços por Atacado–Total da Indústria*, covering all industrial sectors; by *FGV*), *IPA-TRANSF* (*Índice de Preços por Atacado–Transformação*, covering manufacturing sectors; by *FGV*), *IGP-DI* (*Índice Geral de Preços–Disponibilidade Interna*, consumer price index covering domestically produced commodities and services; by *FGV*), and *INPC* (*Índice Nacional de Preços ao Consumidor*, national consumer price index; by *IBGE*).

Some sector-specific wholesale price indices are available for Brazilian manufacturing sectors between 1986 and 1998. The two most natural choices seem to be *IPA-OG* (*Índice de Preços por Atacado–Oferta Global*) and *IPA-DI* (*Índice de Preços por Atacado–Disponibilidade Interna*). Both series are calculated and published by *Fundação Getúlio Vargas (FGV), Rio de Janeiro.* They are wholesale price indices. Brazil disposes of no producer price index for the period 1986-1998. While *IPA-DI* restricts attention to the wholesale of domestically manufactured products, *IPA-OG* includes both imported and domestic goods.

Industry-wide price indices permit deflating all variables in the same manner. As soon as sector-specific indices are indicated, however, different flow variables have to be deflated using different indices. Appropriate choices for different types of flow variables are discussed in the following subsections.

A.4.3 Price indices for output variables

Being wholesale price indices, neither *IPA-OG* nor *IPA-DI* reflect the price level at the sales gate of the manufacturers. Still, these series seem to come close to proper sector-specific output deflators in Brazil. Neither *IPA-OG* nor *IPA-DI* use sector definitions that coincide with the sector classification in *PIA*. Firms in *PIA* are categorized according to *IBGE*'s *nível 100* system (its degree of detail corresponds roughly to three *SIC* digits). Tables in appendix A.7 (p. 243) propose how to make the sectoral classifications conform. There are 62 industrial sectors within *nível 100*.

I apply these price indices to the output related variables `grssales`, `sales`, `difstock`, and `resales` in *PIA*.

A.4.4 Price indices for inputs of intermediate goods

While wholesale price indices may provide adequate series for deflating output, they seem arguably less appropriate for the prices at the firm's gate for purchases. Prices at the input side and at the output side of firms are likely to behave differently in periods of high or volatile inflation. Therefore, I use the national input-output matrices to derive the typical input basket of a firm in a given sector. With this information at hand, sector-specific input price indices are constructed.

The national accounting department at *IBGE* calculates yearly input-output matrices. With the change in the system of national accounts after the 1990 census, however, time-consistent matrices are only available for the years 1990 to 1998, and for 1985. The year 1985 is used to link the 1990 accounting standard to earlier systems. In order to obtain comparable input-output matrices for the entire period 1986-1998, I construct the matrices for 1986 through 1989 as intermediate matrices between the two known matrices for 1985 and 1990. A linear interpolation is applied.

The input-output matrices under the 1990 system are 80×43 matrices—the 80 rows representing the economy-wide sectors at *nível 80* from where the inputs come, and the 43 columns representing the sectors according to *nível 50* to which the inputs go.[5] For the purpose of deflating variables in *PIA*, not quite as many rows and columns (sectors) are needed. Among the 80 rows at *nível 80*, only 52 correspond to industrial sectors. Similarly, among the 43 columns at *nível 50*, only 30 correspond to industrial sectors. I use the reduced 52 by 30 matrices for the calculations to follow. This reduction disregards non-industrial inputs but non-industrial inputs are only a negligible share of total inputs in manufacturing.

For the construction of sector-specific input price indices, only relative weights of those sectors are needed where inputs come from. Due to the form of the input-output matrices, it is the columns which provide these weights. To obtain them, we can express the entry in each cell of the input-output matrix as a fraction of the sum of the entries in the respective column. An example is given below.

$$\mathbf{X} = \begin{pmatrix} 100 & 300 & 0 \\ 100 & 200 & 0 \\ 100 & 500 & 100 \\ 100 & 0 & 0 \end{pmatrix} \rightarrow \mathbf{A} = \begin{pmatrix} .25 & .3 & 0 \\ .25 & .2 & 0 \\ .25 & .5 & 1 \\ .25 & 0 & 0 \end{pmatrix}$$

In general, take the input-output matrix $\mathbf{X}$ and call the entry in row i and column j x_{ij}. I obtain the matrix of weights $\mathbf{A}$ by placing the entry $a_{ij} = x_{ij}/(\sum_i x_{ij})$ in cell (ij) and linearly construct substitutes for the missing input-output matrices between 1986 and 1989. Call every entry in the weights matrix in 1985 a_{ij}^{85} and call every entry in the 1990 weights matrix a_{ij}^{90}. The intermediate weights for the years $t = 86, 87, 88, 89$ are

$$a_{ij}^{t} = a_{ij}^{85} + (t - 85) \cdot \frac{a_{ij}^{90} - a_{ij}^{85}}{5}. \tag{A.3}$$

[5] *Nível 50* is equivalent to *atividade 80* and *atividade 100*. It coincides with the first two digits of both *nível 80* and *nível 100* and roughly corresponds to two *SIC* digits.

This procedure yields weights matrices for 1986 through 1989 whose columns sum to 1 (since $\sum_i(a_{ij}^{90} - a_{ij}^{85}) = 0$ and $\sum_i a_{ij}^{90} = 1$). Their values linearly reflect the change in the input-output structure over the five-year period.[6]

Finally, call the vector of output price indices for month m in year t $\pi_{output}^{m,t}$. I obtain the vector of sector-specific input price indices as

$$\pi_{input}^{m,t} = (\mathbf{A}^t)' \pi_{output}^{m,t}. \tag{A.4}$$

For the deflation of data in *PIA*, I depart from the (wholesale) price indices as described in subsection A.4.3 above. Then the vectors $\pi_{output}^{m,t}$ represent the 62 industrial sectors at *nível 100*. To make these 62 sectors conform to the 52 industrial sectors at *nível 80*, the price indices need to be averaged at *nível 50*, and $\pi_{output}^{m,t}$ is accordingly reduced to 52 rows.[7] The weights matrix $\mathbf{A}^t$ has dimensions 52×30. So, the resulting input price vector $\pi_{input}^{m,t}$ has 30 rows—representing the 30 industrial sectors at *nível 50*.

A.4.5 Price indices for inputs other than intermediate goods

Under inflation, economic variables such as salaries, financial expenditures and rental or leasing rates tend to respond more or less in line with money supply. Accordingly, they are often deflated by economy-wide consumer price indices such as *INPC* or *IGP-DI*. However, in the context of a firm's decision making process, the use of a less general deflator may be more appropriate. For the firm, its decision to substitute between factors of input (capital and labor, say) or between different forms of employing these factors (make or buy or rent) depends on the relative prices of these alternatives, and the relative sales price for final products. Therefore, a more adequate choice may be the use of industry-specific rather than economy-wide price indices. In particular, the use of the *IPA-OG* and *IPA-DI* series for deflating outputs and intermediate goods inputs suggest the use of the industry-wide prices indices within the *IPA-OG* or *IGP-DI* series, too, to deflate the above-mentioned economic variables. The appropriate choice of a deflator for `profit` is less clear. However, since balance sheet profits also serve as an indicator for the management's evaluation of a firm's success and since profits derive from industry-specific activity, the use of indices such as *IPA-OG* or *IGP-DI* may again be most adequate.

I apply these price indices to the variables `wagetop`, `wagewh`, `wagebl`, `asrtimmo`, `aslsimmo`, `fincost`, and `profit` in *PIA*. Depreciation costs `deprec` are treated like total asset retirements `asltot` (see section A.5.4).

[6] The construction of a geometrically evolving series of input-output matrices proves infeasible with common microcomputer capacity. The RAM of a typical personal computer does not suffice to take the fifth root of the 30×30 square matrix $(\mathbf{A}^{85\prime}\mathbf{A}^{85})^{-1}\mathbf{A}^{85\prime}\mathbf{A}^{90}$.

[7] Where ever possible, finer matches between *nível 80* and *nível 100* are chosen.

Table A.11: Price Indices for Gross Investment Flows

Group	Name	Sectors (*nível 80*)[a]
1	`buildings`	(general index)
2	`machinery`	801, 1101
3	`vehicles`	802, 1201, 1301
4	`computers`	1001[b]
5	`other`	1401, 3201
6	`total`	(capital formation weights)

[a]For a list of sectors at *nível 80*, see appendix A.7.4.

[b]Only uses sector 1030 at *nível 100*.

A.4.6 Price indices for gross investment flows

There are six main groups of investment flows in both *PIA velha* and *PIA nova*: (1) buildings, (2) machinery, (3) vehicles, (4) computers, (5) other investment goods, and (6) total investment flows. This section discusses asset purchases in these six categories (gross investment flows). Asset retirements need to be treated differently and are discussed in section A.5.4 below. For the groups (2) through (5), appropriate price indices are constructed using the average of adequate sector-specific (wholesale) price indices. Table A.11 shows the sectors over which the according price indices are averaged. The weights for the averages are obtained from the national capital formation vector for Brazil, which is explained below.[8]

Deflating total gross investment (group (6)) is more intricate. If the national accounts in Brazil provided sector-specific capital formation statistics, investment flows could be deflated by indices similar to the ones constructed for intermediate goods (in subsection A.4.4). However, for the period until 1998 *IBGE* does not break capital formation down into sectors. Instead of a capital formation matrix, *IBGE* only provides a capital formation vector for the economy as a whole, containing the sectors whose output is used for capital formation. I use the (normalized) entries in this capital formation vector as weights for a price index to deflate total gross investment, and as the weights for groups (2) through (5). The capital formation vector is based on the industry classification at *nível 80*. Capital formation vectors between 1986 and 1989 are missing. They are constructed through linear interpolation. Calling an entry in the capital formation vector in 1985 a_{ij}^{85} and an entry in the 1990 vector a_{ij}^{90}, the intermediate entries for the years $t = 86, 87, 88, 89$ result as

$$a_{ij}^{t} = a_{ij}^{85} + (t - 85) \cdot \frac{a_{ij}^{90} - a_{ij}^{85}}{5}.$$

[8]For the purpose of transforming deflators at *nível 100* to deflators at *nível 80*, the finest possible mapping between *nível 80* and *nível 100* is derived through algorithms. Sectors 801 and 802, for instance, are separated and correspond one-to-one to 810 and 820, respectively; it is thus not always necessary to move through *nível 50*.

This procedure yields proper weights for 1986 through 1989, and their values linearly reflect the change in the capital formation structure over the five-year period.

Call the vector of output price indices for month m in year t $\pi^{m,t}_{output}$. Call the vector of weights, derived from the capital formation vector, $\mathbf{a}^t$. I then obtain the economy-wide gross investment flow deflator as

$$\pi^{m,t}_{investment} = (\mathbf{a}^t)'\pi^{m,t}_{output}, \tag{A.5}$$

a scalar. In the case of *PIA*, I depart from the (wholesale) price indices as described in subsection A.4.3 above. Then the vectors $\pi^{m,t}_{output}$ represent the 62 industrial sectors at *nível 100*. To make these 62 sectors conform to the 52 industrial sectors at *nível 80*, the price indices need to be averaged at *nível 50* (or the finest possible mapping above), and $\pi^{m,t}_{output}$ is accordingly reduced to 52 rows. The weights vector $\mathbf{a}^t$ has 52 rows.

I apply the group (2) price index to the variables `acqmasum`, `acqmadom`, `acqmause`, and `acqmafor`. The group (3) price index is applied to `acqveh`, the group (4) index to `acqcomp`, and the group (5) index to `acqother`. The group (6) index seems most appropriate for `acqtot` and possibly `acqbl`. However, I deflate `acqbl` in group (1) with the general price index *IPA-DI* (or *IGP-OG*). Alternatively, a construction price index series could be used.

A.5 Deflating Assets and the Construction of Capital Stock Series

As mentioned in the preceding section A.4, *Legislação Societária* mandates that firms correct the values of their assets in the balance sheet every year. It further requires that they do this correction on the basis of a governmentally administered price index. *PIA* requests that firms report all variables according to this law. The official price index generally tends to understate true inflation. This creates a first bias in the reported stock variables in *PIA*. The bias becomes sizeable over the years. In 1991, the federal government allows firms a once-and-for-all correction of this bias. *Lei n. 8200 de 28-6-91* and the according order *Decreto n. 332 de 4-11-91* to enforce it give all firms the option to revalue their capital stock between January 1991 and December 1991 (Rodrigues, Pereira da Silva and Barros 1992). Firms have strong incentives to revalue their capital stock since they can increase the value of their assets without being taxed for it, and will be allowed to claim the increased depreciation cost in their income statements from 1993 on, thus lowering profits and corporate taxes. *PIA* does not allow to directly observe which firm opts for the correction of the capital stock. These facts make it difficult to construct a capital stock series from balance sheet data. However, there are reasonably precise ways to correct for the two possible biases.

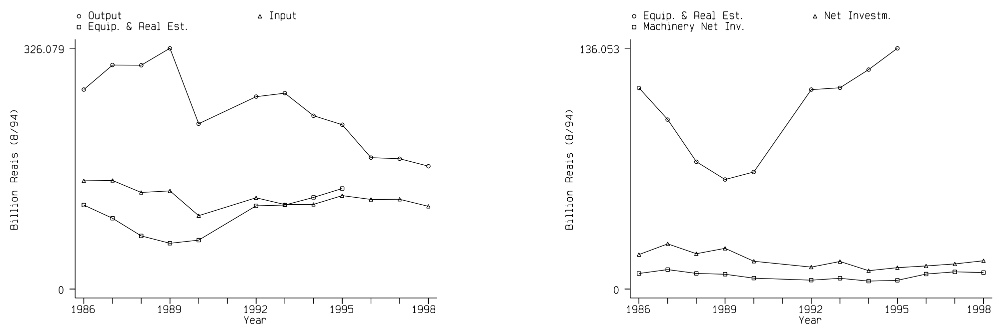
Constructing a capital stock series from net investment flows (using a perpetual inventory method, say), is not safe from these two biases either. The reason is that asset retirements in *PIA* are recorded with the remaining book values at the time of the asset retirement.[9] So, whereas gross investments are properly deflated using price indices as described in subsection A.4.6, the asset retirements counterpart is most likely not deflated correctly with these indices. As a consequence, net investment flows can only be properly inferred when remaining book values are known.

A.5.1 Judging consistent capital stock series

Figure A.2 shows two series of relevant economic variables in *PIA*. While the flow series are deflated as described in the preceding section A.4, the asset series are treated as if none of the aforementioned potential pitfalls existed. The reported year-end values in *PIA* are merely adjusted to a common base month (August 1994). I will subsequently call this series the raw series. Compared with output and value added fluctuations, changes in the capital stock may even seem moderate. The capital stock is measured as the total of ground and premises, machinery, vehicles, and other equipment (`aspimmo`, for *Ativo Imobilizado*). However, when taking the net investment flows both for the *Ativo Imobilizado* (`acqtot` and `asltot`) and just for the machinery part within the *Ativo Imobilizado* (`acqmasum` and `aslmasum`), their fluctuation cannot explain the change in the capital stock—unless there is a negative depreciation rate in 1992. There are mainly two peculiarities about the series. First, the capital stock falls between 1986 and 1988 while net investment flows remain constant. This could be explained by a changing depreciation rate that was higher before the modernization of the capital stock in the late eighties, and possibly by high capacity utilization, wearing the capital stock out. Second, the capital stock jumps in 1992. This is entirely unexplicable with the other data series. Unless there is a huge unobserved jump in investment in 1991 (the missing year in *PIA*), which is unlikely given the general economic situation in Brazil that year, investment flows are at odds with an increase in the stock in 1992. This jump is most likely a consequence of the optional asset revaluation in 1991.

Given the fact that both net investment flow and capital stock series are constructed in *PIA*, we can, in principle, play one against the other until we find two mutually consistent series. An immediate criterion for consistency is, for instance, that the implicit depreciation rate behind the two series must not turn negative in any year. The missing year 1991 makes it difficult but not impossible to design algorithms based on several criteria such as: 'no negative implicit depreciation rate', 'no increases or decreases of more than x per cent in any two consecutive years', and are

[9]Following Brazilian accounting principles, a possible difference between the sales prices for a retired machine and the book value enters the profit or loss account as extra-ordinary revenue of cost.



Data: Unbalanced panel of all firms in *PIA* 1986-1998. Figures are unweighted sums.

Figure A.2: Value added, net investment, and the raw capital series

extended to criteria such as ‘abnormally high implicit depreciation rates only in years with high capacity utilization’, and so forth. It is straightforward to confirm that the data series depicted in figure A.2 violate several reasonable criteria.

In subsections A.5.3 and A.5.4 below, I elaborate a method to measure the potential bias in book values of assets and assets sales, respectively. However, after applying the measures to *PIA*, the resulting capital stock and investment flow series do not meet several other consistency criteria. While the re-valuation jump in 1992 will disappear, intermediate capital stock values in 1989 and 1990 start to violate consistency. Subsection A.5.7 will finally describe a method to correct the capital stock series in a way that satisfies reasonable consistency criteria. To start, the following section briefly describes the series of governmentally imposed price indices by which assets have to be valued.

A.5.2 Governmentally imposed deflators

A recast of the governmentally imposed official price index makes part of almost every Brazilian plan to combat inflation until 1994. Several of these indices deliberately underestimate true inflation. Between January 1986 and December 1994, the combined series of official price indices reports an average annual inflation rate of around 710 per cent. True inflation is about 820 per cent (as measured by *IBGE*'s *INPC*). The according indices since 1964 are:

- *ORTN* (*Obrigação Reajustável do Tesouro Nacional*) in force from October 1964 until January 1989, renamed to *OTN* (*Obrigação do Tesouro Nacional*) under *Plano Cruzado* in 1986

(*Decreto-lei n. 2284/86*). There are two series for the year 1986, one applicable to assets (frozen between March 1986 and February 1987) and the other applicable to asset retirements (continuously adjusted every month).

- *BTN* (*Bônus do Tesouro Nacional*) in force from February 1989 until January 1991 (*Lei n. 7777/89*).
- *FAP* (*Fator de Atualização Patrimonial*) in force for the months February until December 1991 (*Decreto n. 332 de 4-11-91* retroactively).
- *UFIR* (*Unidade Fiscal de Referência*) in force since January 1992. For the period January 1992 through August 1994, daily values are provided (*UFIR Diária*, *Lei n. 8383/91*); beginning-of-month values are generally to be used for deflating monthly figures. For the period September until December 1994, monthly values are provided (*UFIR mensal*, *lei 9069/95* retroactively). Quarterly values of *UFIR* are calculated from January 1995 on, half-year values from January 1996 on, and since January 1997 yearly values (*lei 9069/95*) are provided. (*UFIR* will finally be repealed in October 2000.)

I combine these official price indices to two consistent monthly series of governmentally imposed price indices. Due to a different treatment in 1986, one series has to be applied to assets (`govdefl-asset`), and another series to asset retirements (`govdefl-decap`). The proper links of the indices over time are documented in IOB (2000), for example.

A.5.3 Stock variables

Suppose the capital stock of a firm (or, for the present purpose, one asset position in the balance sheet) is composed of many different single units $i = 1, \ldots, N$. The value of each unit i at the date of purchase $t_0(i)$ is $k_i(t_0(i))$. For simplicity, call it k_i. This unit i wears out and depreciates, and its value needs to be adjusted for inflation. A firm thus calculates the total value of its capital stock (or a position in its balance sheet) at time t using a formula like

$$K_t = \sum_{i=1}^{N} \frac{\pi_t}{\pi_{t_0(i)}} \cdot \delta_{t,i} \cdot k_i, \tag{A.6}$$

where total depreciation of each unit at time t is given by $\delta_{t,i} \equiv \prod_{s=t_0(i)}^{t} \delta_s$. The main issue is the application of an adequate price index π_s at times $s = t$ and $s = t_0(i)$.

By law, firms are forced to use the governmentally administered price index (*ORTN* through *UFIR*), which understates inflation. Call this price index π_s^{otn}. It underlies the asset value $\widetilde{K}_t$

Table A.12: Lifetime of Assets by Brazilian Accounting Standards

Group	Name	Lifetime z in years[a]	Deprec. Rate δ down to 10%	Deprec. Rate δ down to 5%
1	`buildings`	25	.088	.113
2	`machinery`	10	.206	.259
3	`vehicles`	5	.369	.451
4	`computers`	4	.438	.527

[a] These rule-of-thumb lifetimes apply particularly to the electrical and electronics industry. They are not expected to differ in similar sectors, but may differ to some degree for industries such as chemicals and pharmaceuticals.

reported in *PIA*. So, the true value of the capital stock exceeds the reported value of the capital stock by a factor of

$$\frac{K_t}{\widetilde{K}_t} = \frac{\sum_{i=1}^{N} \frac{\pi_t}{\pi_{t_0(i)}} \delta_{t,i}\, k_i}{\sum_{i=1}^{N} \frac{\pi_t^{otn}}{\pi_{t_0(i)}^{otn}} \delta_{t,i}\, k_i} = \frac{\pi_t}{\pi_t^{otn}} \frac{\sum_{i=1}^{N} \frac{1}{\pi_{t_0(i)}} \delta_{t,i}\, k_i}{\sum_{i=1}^{N} \frac{1}{\pi_{t_0(i)}{}^{otn}} \delta_{t,i}\, k_i}. \tag{A.7}$$

This factor is equal to unity if both the true and the imposed price index series are the same at all times. If, on the other side, the imposed price index falls short of the true price index in every period t except for one initial period $t_0(i)$ at which, suppose, all assets have been purchased, the factor takes a value of π_t/π_t^{otn}.

The remaining task is to find a reasonable value for factor (A.7) that is not comprised by too strong assumptions. The procedure proposed here is based on three assumptions. Assumption (a): All firms apply a linear depreciation method. This method equally distributes the initial value of the asset over the years of its likely use. In practice, a large number of Brazilian firms applies this linear depreciation for most assets. Assumption (b): The average life-time of an asset is given by the third column in table A.12. Assumption (c): Every position in the balance sheet is composed of a number of units N. It is equally likely that any one of these N units is acquired in January, February, or any other month of a given year.

The alleged lifetimes under assumption (b) may seem low when contrasted with comparable geometric depreciation rates. For a comparison, the fourth and fifth column in table A.12 list the corresponding values for an annual depreciation rate δ if geometric depreciation is applied and if the asset still has 10% or 5% of its value at the end of the lifetime. To my knowledge, there are no precise estimates for capital depreciation in Brazil as of today. Even if such measures exist, it seems more adequate to neglect them and to use the typical choices of asset lifetimes made by Brazilian accountants, instead. After all, we are interested in correcting typically employed accounting methods retroactively.

It is difficult to judge how strong (and possibly wrong) assumption (c) is in practice. In general, the less frequently an asset is purchased or improved, or the lower its turnover, the more misleading is assumption (c). However, the lifetime of an asset may not be very indicative of whether assumption (c) is too strong for the asset or not. Even though buildings and machinery are purchased at distant intervals, they are also continuously renewed through appropriate services, overhauls, and renovation work. By Brazilian accounting standards, this increases their value again. Similarly, if a firm is large it is likely that the capital stock in different units of operation is replaced with new equipment at different times. This smoothes out possible errors from assumption (c).

Assumption (a) implies that the geometric depreciation term $\delta_{t,i}k_i$ in (A.6) and (A.7) needs to be replaced appropriately. Take computers as an example. Every computer, no matter when purchased, is supposed to remain in use for four years by assumption (a). Assumptions (b) and (c) then imply, for any given point in time, that one quarter of the computer stock is between 37 and 48 months old, one quarter between 25 and 36 months, and so fourth. Or, more precisely, the oldest 48^{th} of the computer stock has only one 48^{th} of its lifetime to survive, the second-oldest 48^{th} will remain in use for two more 48^{th}s of its lifetime, and so on. So, we can assign a weight to every 48^{th} of the capital stock, the oldest receiving a weight of 1 and the youngest a weight of 48. These weights sum up to $(1+48)\cdot(48/2) = 1,176$.

More generally, we can call each of these weights σ_m, representing the bundle of units $(i(m))$ that are acquired in month m. For a capital good with a lifetime of z years, there are $12z$ months of use. Thus,

$$\sigma_m \equiv \frac{m}{(1+12z)6z},$$

where $m = 1, \ldots, 12z$. Hence, the formula in (A.7) can be rewritten to

$$\frac{K_t}{\widetilde{K}_t} = \frac{\sum_{m=1}^{12z} \frac{\pi_{12z}}{\pi_m} \sigma_m}{\sum_{m=1}^{12z} \frac{\pi^{otn}_{12z}}{\pi^{otn}_m} \sigma_m} = \frac{\sum_{m=1}^{12z} \frac{\pi_{12z}}{\pi_m} m}{\sum_{m=1}^{12z} \frac{\pi^{otn}_{12z}}{\pi^{otn}_m} m}. \tag{A.8}$$

In this notation, $12z$ represents December of the respective year in *PIA*. Formula (A.8) is a backward looking expression and specific to each type of capital good (it depends on z, which differs for different capital goods). In this formula, the backward horizon lies the further in the past the longer a typical unit of capital is supposed to be in use.

This method is applied to `aspimmo`, `aspdefer`, and `aspmasum`. Some liability variables and some further asset variables are also deflated with this method.

A.5.4 Asset retirements

Asset retirements in *PIA* are adjusted in a manner similar to the correction of capital stock values. Call the asset retirements in a given year t S_t. Then, similar to equation (A.6), we can write

$$S_t = \sum_{j=1}^{M} \frac{\pi_t}{\pi_{t_0(j)}} \cdot \delta_{t,j} \cdot k_j^s, \tag{A.9}$$

where k_j^s now denotes capital good j, which is acquired at $t_0(j)$ and is being sold at time t. Call the observed figure of asset retirements in *PIA* $\widetilde{S}_t$.

Under the same assumptions (a) through (c) as made for capital stock variables in subsection A.5.3 above, the adjustment factor for asset retirements becomes

$$\frac{S_t}{\widetilde{S}_t} = \frac{\sum_{m=1}^{12z} \frac{\pi_{12z}}{\pi_m} m}{\sum_{m=1}^{12z} \frac{\pi_{12z}^{otn}}{\pi_m^{otn}} m}. \tag{A.10}$$

Assumption (c) is a little more restrictive in this context. Assumption (c) states that each asset sold today is equally likely to have been acquired in any preceding month. It seems more probable, however, that old capital goods are sold more often than recently purchased ones. There are possible adjustments to formula (A.10) such as replacing the factor m in both the numerator and the denominator by a less steeply increasing function of m. However, these adjustments seem as arbitrary as leaving the formula in this form. However, the factor (A.10) can be biased. The direction of the bias depends on the differences in inflation between more recent periods and more distant periods in the past. If inflation rates are very high in the distant past but lower lately, factor (A.10) tends to be too low and to understate the necessary correction. Similarly, if inflation rates are comparatively low in the distant past and higher lately, factor (A.10) tends to be too high and to overestimate the necessary correction.

This method is applied to variables `asltot`, `aslbl`, `aslmasum`, `aslveh`, `aslother`, `aslcomp`, and `deprec`, the annual total depreciation cost.

A.5.5 Optional revaluation of assets in 1991

A change in Brazilian tax law in 1991 allows firms to revalue their assets in the balance sheet and to correct the bias that the governmentally imposed price index have caused over the years. It is highly likely that many firms opt for this possibility in 1991. The value increase is not taxed but higher depreciation rates in the future will reduce taxable profits. To my knowledge, there are no statistics that would allow to infer which firms choose to revalue in 1991 and which do not. After 1991, the government requires again that firms adjust their asset values on the basis of the official price index, which continues to understate inflation.

So, the observed asset values in *PIA* between 1992 and 1994 suffer from a down-ward bias again. However, this bias is weakened for all firms that choose to revalue in 1991. Assets of revaluing firms have the right value in 1991 and suffer from a lack of correction only thereafter. Hence, formulas (A.8) and (A.10) need to be modified in the case of those firms. For any month prior to January 1992, the ratios $\pi_t/\pi_{t_0(i)}$ in the numerator and denominator need to be replaced. I replace them with the value that they take in December 1991, the date at which revaluation allows to make all asset values precise. In the case of non-revaluing firms, on the other side, assets and asset retirements are still correctly valued by formulas (A.8) and (A.10).

How, then, can one arrive at proper asset values after 1991? There are two types of assumptions one can make. Assumption (i): Almost every firm in *PIA* revalues in 1991, and the few firms that do not revalue have little to correct. *PIA velha*'s firms are medium-sized to large manufacturers, and certainly have the resources to undertake a revaluation of their assets. In addition, most Brazilian firms keep more than one book—one of them being for internal use and more accurate. So, the necessary information is already available to these firms. Moreover, the prospect of tax savings is a strong incentive for them to show the revaluation in their balance sheet. Finally, the few firms in *PIA* that choose not to revalue in 1991 do not expect big tax savings, that is, they expect only small corrections in value. Thus, under this assumption, we commit hardly any error by considering all firms as having effectively revalued. It also seems a likely scenario, and thus safe to adopt.

Assumption (ii): Only a share of firms in *PIA* revalues, accounting cost do not outweigh tax savings. This seems the less likely scenario. Yet, one may find it worthwhile to program algorithms that identify potential 'revaluers.' For this purpose, the 1990 capital stock could be extrapolated, and average net investment flows could be used to identify firms whose jump in the capital stock in 1992 relative to the extrapolated value implies too low a depreciation rate. However, there are two drawbacks to this method. First, the year 1991 is missing in *PIA* and any extrapolation over two years, from 1990 to 1992, will be vague. Second, the threshold for 'too high a jump' (or 'too low an implicit depreciation rate') is arbitrary. When one then corrects the capital stock of the jumpers, one likely reduces the correction factor (A.8) by the size of the threshold. So, the method risks to become circular. Thus, who considers assumption (ii) the more likely scenario should be cautioned from using capital stock series in *PIA* at all. There may not be useful statistical means to correct for the problems arising from the assumption.

Table A.13: Correction Factors for Asset Figures

	Group	Lifetime	*IPA-OG* (at *nível 80*)[a]	*IPA-DI*[b]
1	`buildings`	25	(general index)	ipadi (or igpdi)
2	`machinery`	10	801, 1101	maq
3	`vehicles`	5	802, 1201, 1301	veiculos[c]
4	`computers`	4	1001[d]	bcd-ud
5	`other`	6[e]	1401, 3201	ipadi
6	`total`	14[f]	(capital formation)	ipadi

[a]For a list of sectors at *nível 80*, see appendix A.7.4. Weights according to annual capital formation vector.

[b]For the definition of abbreviations see appendix A.7.3.

[c]Series for *trans* are only available after 1986 and thus not applicable.

[d]Only uses sector 1030 at *nível 100*.

[e]Hypothesized value.

[f]Inferred from typical capital stock composition in *PIA*.

A.5.6 Correction factors for asset figures

The key terms in both formula (A.8) and (A.10) are the ratios π_{12z}/π_m and $\pi_{12z}^{otn}/\pi_m^{otn}$. They are correction factors to undo valuation errors retroactively. In a notation closer to the initial one, they could also be written as π_t/π_m and π_t^{otn}/π_m^{otn}. Here t corresponds to December of the respective year in *PIA* $(86, \ldots, 98)$, and m denotes any month in the $4, 5, 10$, or 25 years preceding t. The correction method proceeds in two steps.

First, for every year in *PIA* the correction factors π_t/π_m and π_t^{otn}/π_m^{otn} are derived. There are six groups of capital goods for which they need to be constructed as shown in table A.11. Table A.12 lists the four groups for which accounting assumptions on lifetime are typically made.[10] I use average price indices in the case of groups (2) through (5) as indicated in table A.13, while a construction price index or a general price index such as *IGP-DI* or *IPA-DI* seem most appropriate for buildings (1).

The underlying price indices for buildings would need to range back until 1961. However, even the governmentally imposed price index *ORTN* only dates back to October 1964. For all present purposes January 1965 is used as first available month. The ratios π_t^{otn}/π_m^{otn} are set equal to the oldest available observation before that date. Similarly, the ratios π_t/π_m are set to the January 1969 value for years before 1969 when *IPA-OG* and *IPA-DI* are used. Finally, the price index *INPC* (possibly useful for buildings) is only calculated since March 1979. For the preceding months and years, it seems most adequate to use the historic price index series *IGPC-MTb* (*Índice Geral de*

[10]As far as pure manufacturing firms are concerned, soil is not exhausted and does not need to be depreciated. In the case of mineral or metal extraction, a further series for `ground` would need to be constructed that applies a weighting scheme different from formulae (A.8) and (A.10) to account for the loss in value due to extraction.

Preços ao Consumidor-Ministério do Trabalho), a national consumer price index provided by the Brazilian federal labor ministry at the time (IBGE 1990).

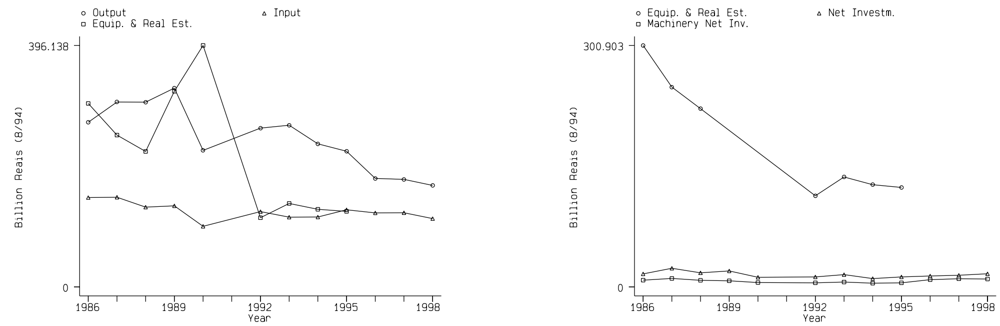
The lifetime for other assets is hypothesized and the average lifetime for total assets is inferred from a typical capital stock composition in *PIA*, given the accounting lifetimes for the preceding categories in table A.13. I make the following back-of-the-envelope calculation for that purpose.

	Gr. Inv. (86-98)	Cap. Stock (86-90)	Turnover	Lifetime
buildings	25.4 %	34.7 %	0.73	25 years
machinery	48.7 %	48.7 %	1.00	10 years
vehicles	4.1 %	2.4 %	1.75	5 years
computers	1.5 %	0.7 %	2.25	4 years
other	20.4 %	13.7 %	1.49	→ *6 years*
total	→ *12.7 years*	→ *14.5 years*		

While the investment flows are known between 1986 and 1998 for all types of capital goods, stocks are only known from the first part of *PIA velha* between 1986 and 1990. The ratios of flows to stocks indicate that 'other capital goods' exhibit an intermediate turnover between machinery and vehicles. So, six years are hypothesized as their average lifetime. With these numbers at hand, the average lifetime of the total capital stock is between 12 and 15 years. 14 years are used subsequently. The reason for using a value closer to the upper bound is that the book values of land is generally not depreciated. As land is part of the total assets, too, the average lifetime of total assets might be understated when excluding land from the calculation.

The accordingly corrected end-of-year values are still current values. They need to be taken to some common base year. This is done by applying the respective indices in table A.13 again. In order to arrive at year-end values, the January and December price indices around the respective year-end are averaged if they are mid-month indices.

Putting this procedure to work yields the capital stock and net investment series shown in figure A.3. There are two new peculiarities about the series. First, the correction factors for the years 1989 and 1990 become extremely large, pushing the capital stock in these years even further up then before. There is no movement in either investment flows or output that could justify this jump. The years 1989 and 1990 are two years of extremely high inflation and economic uncertainty. While the first fact pushes the capital stock series up, the second suggests that the method may be particularly wrong in these years. In periods of high uncertainty, turnover of capital goods is low, gross investment will be low, and there will be few asset retirements. The method gives a high weight to recently purchased assets, however, since they are the least depreciated while considered equally likely to enter the capital stock now as decades ago. This boosts the correction factors in



Data: Unbalanced panel of all firms in *PIA* 1986-1998. Figures are unweighted sums.

Figure A.3: Value added, net investment, and a preliminary capital series

1989 and 1990 (up to factors of 6, depending on the hypothesized lifetime and price index). The graph on the right in figure A.3 therefore ignores these outlier years.

Second, the capital stock is continuously falling through 1986 until 1990, while net investment both in sum and for machinery hardly responds (the method has a levelling effect on net investment through its adjustment of asset retirements). The implied annual depreciation rate between 1986 and 1990 is very high (25 per cent in 1987 and 18 per cent in 1988), while it attains reasonable levels (of about 14 and 12 per cent) after 1992.[11] This may seem unreasonable; it would be ruled out by a criterion on implicit depreciation such as: 'Reject a series if implicit depreciation reaches 20 per cent or more in any year'. Note that the capital stock in this definition includes buildings, which depreciate little even under high capacity utilization.

A.5.7 A method satisfying consistency criteria

While seemingly compelling from a theoretical point of view, the correction method is not likely to pass reasonable criteria on implicit depreciation. The method has the advantage, however, to provide a theoretically well-grounded correction factor for the effects of the optional revaluation in 1991. In a word, it seems reasonable to keep the uncorrected values for the capital stock before 1990. In addition, the capital stock figures after 1991 are readjusted by an appropriate factor to make them comparable to 1990 values. There are several arguments in favor of this procedure.

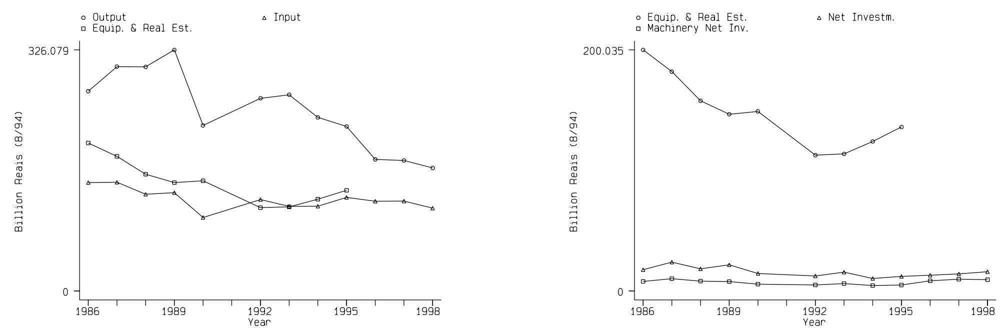
First, the worsening picture after applying the correction factors lends support to the

[11] Since the capital stock K_t evolves according to the relationship $K_{t+1} = J_{t+1} + (1 - \delta_t)K_t$ under net investment I_{t+1} and depreciation $\delta_t K_t$, the implicit depreciation rate is inferred as $\delta_t = [I_{t+1} - (K_{t+1} - K_t)]/K_t$ in every year.

hypothesis that the values in *PIA* are not that far off the mark after all. Second, when the correction factor method does a good job—before 1988 and after 1992—, the rates of change in the capital stock exhibit the same tendencies as the raw series. Between 1986 and 1988, the partially corrected capital stock falls by 26 per cent, while it falls by 37 per cent in the raw series. From 1992 until 1995, the capital stock in the partially corrected series increases by 9 percent, while the raw capital stock goes up by 21 per cent. So, apart from raising the levels in every year, the partial correction method has a smoothing effect on the series. Except for the hyper-inflation years 1989 and 1990, it resembles the movements in the raw series. The partial correction method confirms the pattern of changes in the raw series, albeit diverging to a certain degree in the absolute figures.

Third, it is highly likely that firms apply a monetary correction to their assets that preserves possibly much of the real asset value. While the governmentally imposed price index forces them to undervalue their assets, firms have strong incentives to exploit ways to keep their book values as close to real values as possible. Income taxation makes this a strictly preferable strategy. Since losses from monetary correction cannot be claimed in full—the governmentally imposed price index supposes that there is no such monetary loss—, firms would forego the profit reducing effect of the correct depreciation of their assets and pay unduly high taxes. Consequently, firms have strong incentives to keep assets from losing book value, to keep depreciation costs possibly high and close to the real depreciation costs, and to reduce their reported profits, that is taxable income, by these depreciation costs in subsequent years in order to save taxes. The only way to achieve this is to pencil in book values of the assets as close to real values as possible. So, the main problem of the series does not seem to be the continuous undervaluation of assets, which to avoid firms had strong incentives. The main problem rather seems to be the optional revaluation in 1991 because firms had incentives to report as high a revaluation as they could.

Fourth, the partial correction method provides a theoretically sound basis for the correction of the revaluation in 1991, the major problem in the raw series. What, then, is the right factor to adjust capital stock figures in 1992 and thereafter to figures before? Annual correction factors are calculated under assumption (i) (in the preceding section), which states that almost all firms revalue and that those that don't have negligibly little to change. The ratio between the correction factor for 1992 and the correction factor of each preceding year shows how far the capital stock in the preceding year should be elevated to make the series conform. In other words, the ratio between the correction factor for 1992 and the correction factor of each preceding year measures the degree of revaluation that firms could reasonably claim to be justified in front of the tax authorities. The year 1992 appears to be the year in *PIA* that contains arguably the least error—above all, it immediately follows the revaluation year 1991. Since the correction factors for 1989 and 1990 have proven to be



Data: Unbalanced panel of all firms in *PIA* 1986-1998. Figures are unweighted sums.

Figure A.4: Value added, net investment, and the corrected capital series

out of reasonable range (due to the underrepresented hyper-inflation in the official price index and firms' strong incentives to avoid applying the official index in full), the correction factors for 1986 through 1988 seem the most appropriate base for comparison to 1992. Dividing the 1992 factor by the mean of the factors between 1986 and 1988 yields a ratio of about 2.04 in the case of *IPA-DI* and of about 2.19 in the case of *IGP-DI*, for instance. This indicates the average difference between a wrongly priced capital good's value in 1986 and its equivalent in 1992. Simply multiplying the capital stock figures between 1986 and 1988 by this ratio would boost high values even higher. I therefore chose to multiply the average capital stock in 1986 and 1990 by the factor, to then subtract the average of the raw figures between 1986 and 1990, and to add the so obtained absolute difference in levels to all raw capital stock figures between 1986 and 1990, for every firm. This procedure shifts the left arm of the series to the North, rather than turning it around the midpoint in 1988. This accounts for the fact that the capital stocks in 1989 and 1990 should get a stronger push than the earlier ones since inflation is particularly high in 1989 and 90. Figure A.4 shows the resulting series in the aggregate.

This procedure yields a declining capital stock over the period from 1986 until 1992. The relative decline that occurs between 1986 and 1988 continues from 1990 to 1992. Substantial political and economic uncertainty marks the period. The relatively levelled net investment flow series implies a substantially higher depreciation rate between 1986 and 1990 (18 per cent on average) than between 1992 and 1995 (4 per cent on average). Improved quality of the capital goods and lower utilization of installed capacity may contribute to this. They are unlikely to explain all the difference so that valuation problems in the capital stock series may in fact remain. The evolution of the series is roughly supported by the reported annual depreciation cost in *PIA* (`deprec`)—a variable measured as well or as badly as annual asset retirements. If one proxies the annual depreciation rate with the ratio between this variable (`deprec`) and the initial capital stock (`aspimmo` at the beginning of the year), the average depreciation rate over all sectors and regions would be 7 per cent between 1986 and 1990, and 5 per cent between 1992 and 1995. Clearly, this decline is less pronounced than the series in figure A.4 suggests. Its direction, however, seems to confirm the overall picture.

Depending on the exact criteria one wants to impose for mutually consistent investment and capital stock figures, the resulting series will differ. It seems likely, however, that they roughly resemble the picture of the series in figure A.4. The capital stock series ends in 1995. The completion of the capital stock series until 1998 in the following subsection will show that a continued capital accumulation throughout the 1990s compensates for the reduction of the capital stock during the late 1980s.

A.5.8 Connecting capital stock series between *PIA velha* and *PIA nova*

PIA nova only offers investment flow variables, and no information on levels. So, to extend the capital stock series beyond 1995, a variant of the perpetual inventory method needs to be applied in one form or another. Following the insights from subsection A.5.7 I use the deflated values of the raw series after 1995 (and adjust the figures before 1991).[12]

Net investment flows result as the difference between gross investment and asset retirements. To derive capital stock figures for 1996 and beyond, an assumption about the likely depreciation rate needs to be made. Consistency suggests to use either values from table A.12, or to apply an average of implicit depreciation between 1986 and 1995. I use an imputation procedure, described in detail in subsection A.5.9. Taken together, these steps allow to construct consistent series of the capital stock and related variables for the firms in *PIA* between 1986 and 1998.

Finally, many firms rent or lease both buildings and equipment. To complete the estimate of a capital stock series, the capitalized value of these rental and leasing rates has to be added in. *PIA* provides two variables, `asrtimmo` and `aslsimmo`, that contain information on rented assets. However, these variables do not allow to distinguish between types of capital goods. So, it will be necessary to make assumptions on their separation if one wants to incorporate rented assets at lower than the aggregate level of the capital stock (*ativo imobilizado*). I include them in the structures variable in chapter 3.

[12] This is further supported by the fact that the earlier understating of inflation now sometimes turns into an overstating. The freezing of *UFIR* over longer periods of time makes the correction factor drop below unity for short-lived goods after 1996.

A.5.9 Total capital: Equipment and structures

The closest variable to the total capital stock in *PIA* is *ativo imobilizado* (`aspimmo`). It embraces everything from real estate and buildings, to equipment, vehicles, and computers. However, no information on capital utilization rates is available. I infer a series for the capital stocks from the data using a perpetual inventory method. I choose this method mostly because it relates best to the afore-mentioned accounting and correction principles that determine the observed balance sheet figures.

Over the course of the years, *PIA* questionnaires are reduced and only investment flows become available in later years while several variables on stocks of capital goods were available before. In addition, rental and leasing cost are only reported as totals so that the rental of subgroups of capital goods cannot be inferred directly. Therefore, the capital stock is divided into three parts for the study in chapter 3: Domestic equipment, foreign equipment, and the remaining parts of the total capital stock (corresponding to *ativo imobilizado* in the balance sheet, plus the present discounted value of the rental stream, less equipment stock). The underlying hypothesis is that rental and leasing is mostly used for buildings and vehicles, and less for equipment.

The following three-step procedure yields a coherent capital stock series for each individual firm. While the underlying depreciation rates are imputed (through linear regression and prediction), the capital stock figures are inferred from the according accounting identity $\overline{K}^{tot}_{t,i} = (1 - \hat{\delta}^{tot}_{t,i})\underline{K}^{tot}_{t,i} + I^{tot}_{t,i}$ for every firm i—a perpetual inventory method. The notation here reflects the timing of the observed balance sheet figures. The beginning-of-year capital stock $\underline{K}^{tot}_{t,i}$ in year t equals the end-of-year capital stock $\overline{K}^{tot}_{t-1}$ of the preceding year.

Step 1: Since no survey is conducted in 1991, the initial *total* capital stock for 1992 is missing. Given an estimate of the depreciation rate, $\hat{\delta}^{tot}_{92,i}$, the initial capital stock in 1992 results as $\underline{K}_{92,i} = (\overline{K}_{92,i} - I_{92,i})/(1 - \hat{\delta}^{tot}_{92,i})$. The firm-specific depreciation rate for 1992 is imputed in two stages: First, a firm-specific depreciation rate $\hat{\delta}^{tot}_{t,i}$ is calculated for every firm and year (86-90, and 93-95) as the ratio between the reported total depreciation cost and the initial total capital stock: $\hat{\delta}^{tot}_{t,i} = D^{tot}_{t,i}/\underline{K}_{t,i}$. Total depreciation cost $D^{tot}_{t,i}$ is an observed variable in *PIA*. Second, regressing this firm and year-specific depreciation rate on a constant and on total depreciation cost allows one to predict the missing firm-specific depreciation rate for 1992. If depreciation cost is missing in 1992 or the regression has too few observations, the predicted sector and region wide depreciation rate $\sum_{i\in(\mathbb{S}\cap\mathbb{R})}^{N} \hat{\delta}^{tot}_{t,i}/N$ is used instead.

Step 2: *PIA* contains no total capital stock figures after 1995. The end-of-year capital stock figures from 1996 until 1998 are inferred as $\overline{K}^{tot}_{t,i} = (1 - \tilde{\delta}^{tot}_{i})\underline{K}^{tot}_{t,i} + I^{tot}_{t,i}$, where $\tilde{\delta}^{tot}_{i}$ is calculated as the firm-specific average between 1992 and 1995: $\tilde{\delta}^{tot}_{i} = \sum_{s=92}^{95} \hat{\delta}^{tot}_{s,i}/4$. Since a structural break may occur between 1990 and 1992, depreciation rates in earlier years are not included at this stage.

Step 3: Firms rent and lease more assets after 1992. In addition, smaller firms rent a larger share of their capital stock. In order to prevent a bias from the higher renting and leasing activity after 1992 and among smaller firms, capital stock equivalents to the rental rates are constructed and added to the proprietary capital stock. Brazil does not dispose of data on rental rates for a firm's typical capital stock.

So, the following procedure is adopted to infer rental rates. Rental and leasing rates must compensate for the user cost of capital, that is for both foregone real interest and depreciation. In equilibrium, the annual rental rate in year t, d_t, must equal the annualized monthly real interest rate in year t plus the typical annual depreciation rate at firm i: $d_t = r_t + \hat{\delta}_{i,t}$. The real interest rate is calculated as the monthly interest rate on a savings account (*poupança*). Researchers regard the monthly savings account interest rate as a good indicator of opportunity cost for investments in Brazil, especially since risk-adjusted yields of assets fluctuate considerably. A consistent savings account interest rate series (*Caderneta de Poupança - Rendimento Mensal*) is available from *Associação Nacional das Instituições do Mercado Aberto* through *Fundação Getúlio Vargas*, Rio de Janeiro (*FGV Dados*). The monthly nominal interest rate is purged of monthly inflation using the national consumer price index *INPC*, and then annualized. The years 1989 and 1990 are disregarded as they are characterized by unexpectedly high inflation, resulting in negative real interest rates of as low as -25%. The rental rates for buildings and equipment cannot have been based on such expectations so that these interest rates are discarded. Instead, for the years 1986 through 1990, the average real interest rate between 1986 and 1988 is used (5.3 percent). Similarly, for the periods 1992 until 1995, and 1996 until 1998, the according four and three-year averages are used (10.3 and 10.0 percent, respectively). The annual depreciation rates are calculated for every firm using the method in step 1. They are then averaged, for each firm, in the same three subperiods to remove fluctuations which are unlikely to have been the basis for rental rates. The rented capital stock then results as $\overline{K}_t^{rent} = D_{i,t}/(\bar{r}_t + \bar{\hat{\delta}}_{i,t})$, where $D_{i,t}$ denotes firm i's rental and leasing expenditure in year t, and $\bar{r}_t$ and $\bar{\hat{\delta}}_{i,t}$ the according period-averages of the real interest rate and the depreciation rate.

Wherever possible, missing values in *PIA*'s capital stock figures are imputed as $\overline{K}_{t,i} = (1 - \hat{\delta}_{t,i})\underline{K}_{t,i} + I_{t,i}$, using an estimate of the depreciation rate as in step 1. *PIA* does not distinguish between missing and zero-value observations prior to 1996. For these early years, missing or zero-value stock observations are assumed to be missing values in fact, whereas missing or zero-value figures for investment flows are considered to be zero if and only if investment flows in similar or related variables are observed. For example, if equipment acquisitions are not observed while equipment retirements are, the missing or zero-value entry is treated as zero. It is left missing if, for instance, total investment flows are observed but no flows related to equipment. Alternatively, I tried direct imputation (regression and prediction) methods for capital stock values. The resulting series were highly volatile and produced a considerable share of unreasonable outliers. Therefore, the mixture of imputed depreciation and inferred stock values seems preferable.

A.5.10 Domestic and foreign equipment

The following five-step procedure yields a coherent equipment stock series.

Step 1: Since no survey is conducted in 1991, the initial *total* capital stock for 1992 is missing. The results from step 1 above are reused (appendix A.5.9).

Step 2: Beginning and end-of-year *equipment* stock figures are available between 1986 and 1990, but not thereafter, and the year 1991 is missing. The initial equipment stock in 1992 is inserted using the average share of equipment in total capital in the beginning of all preceding years 1986 through 1991 (the beginning-of-year value is recorded for 1986, and the 1991 value is inferred from the 1990 end-of-year value): $\hat{\phi}_{92,i} = \sum_{s=86}^{91}(\underline{K}_{s,i}^{mach}/\underline{K}_{s,i}^{tot})/6$. Then, $\underline{K}_{92,i}^{mach} = \hat{\phi}_{92,i}\underline{K}_{92,i}^{tot}$. If the firm is the legal or economic successor of another firm and emerges either in 1991 or 1992, the according ratio of the predecessor firm is used. If a firm is new born or a firm-specific estimate for $\hat{\phi}_{92,i}$ is missing for some other reason, the average of the sector and region is used $(\sum_{s\in\{86,\ldots,90\},i\in(\mathbb{S}\cap\mathbb{R})}^{N}(\underline{K}_{s,i}^{mach}/\underline{K}_{s,i}^{tot})/6N)$. If a firm is created in a year after 1992 by some parent firm, its parent's capital structure is copied. If a greenfield creation emerges after 1992, the typical capital composition in the firm's sector and region is imposed as starting structure.

Step 3: The end-of-year equipment stock between 1992 and 1998 is no longer reported in *PIA*. These values are inferred from the accounting relation $\overline{K}_{t,i}^{mach} = (1-\hat{\delta}_{t,i}^{mach})\underline{K}_{t,i}^{mach} + I_{t,i}^{mach}$, starting in 1992 and moving forward to 1998. When an investment flow is missing in an intermediate year, the average of the equipment flow in two neighboring years is used, weighted by the according total flow figure, in order to preserve subsequent observations. An estimate of the firm and year-specific equipment depreciation rate $\hat{\delta}_{t,i}^{mach}$ is derived applying the following procedure: First, total depreciation rates for every firm and year are computed as in step 1 in the previous subsection A.5.9, using the total depreciation cost reported in *PIA*. Second, since no explicit equipment depreciation cost figure is available in *PIA*, an estimate of the average lifetime ratio between equipment and the total capital stock is obtained. In steady state (and the years 1986 through 1989 are assumed to come close to a steady state), the ratio between the average lifetime of equipment and total capital stock must be equal to the inverse of the ratio between the depreciation rates for equipment and total capital stock. Also, the ratio of average lifetimes can be approximated by average turnover:

$$\frac{\hat{\delta}^{mach}_{t,i}}{\hat{\delta}^{tot}_{t,i}} = \frac{\text{Avg. Lifetime Total Capital}_{(t,i)}}{\text{Avg. Lifetime Equipment}_{(t,i)}} \approx \frac{\frac{I^{mach}_{t,i}}{(\underline{K}^{mach}_{t,i}+\overline{K}^{mach}_{t,i})/2}}{\frac{I^{tot}_{t,i}}{(\underline{K}^{tot}_{t,i}+\overline{K}^{tot}_{t,i})/2}}, \tag{A.11}$$

where average turnover is defined as the annual gross flow divided by the annual average stock. Note that in steady state annual gross investment just replaces depreciated capital $I_t = \delta_t \underline{K}_t$. (Alternatively, the implicit equipment deprecation in the years 1986 through 1990 is calculated as: $\hat{\delta}^{mach}_t(1+(I_t - \bar{K}_s)/\underline{K}_s)$ but figures are found to be too erratic to base further derivations on them.) In *PIA*, the lifetime ratio (A.11) fluctuates strongly across regions and sectors but is fairly stable over the years. On average, it amounts to 1.37. That is, the lifetime of equipment is about 37 percent shorter than that of an average capital good in steady state. Since buildings and real estate enter the total capital stock but depreciate little, this figure seems reasonable. In addition, Brazilian accounting rules of thumb take ten years as the average lifetime of equipment, 25 years for buildings and between four and six years for cars, computers, and the like; this yields an average of roughly 14 years of life for the average total capital stock of a typical Brazilian firm—the ratio of 14 by 10 is close to the figure estimated here. Since it seems more plausible to assume that the industry as a whole found itself in steady state than to assume that every single sector is in steady state, this overall ratio of 1.37 is applied to all sectors. The firm and year-specific equipment depreciation rates are set to $\hat{\delta}^{mach}_{t,i} = 1.37 \cdot \hat{\delta}^{tot}_{t,i}$, where $\hat{\delta}^{tot}_{t,i}$ is the same as in step 1. It is likely that most of the fluctuations in the depreciation cost for a firm come from equipment and short-lived capital goods, rather than from ground and premises. So, observed fluctuations in the overall depreciation rate should be carried through to equipment depreciation. The present method does that.

Step 4: As regards foreign equipment, only acquisitions are observed in *PIA*. They need to be used to infer stock values over the sampling period. Since industry is closest to a steady state in the mid eighties, the following method tries to infer a likely foreign equipment stock in the earliest possible year and to depart from this estimate subsequently. Firms in *PIA* are conveniently split into two groups: (a) Firms born in 1985 or before, and (b) firms born in 1986 or during the sampling period. Turn to group (a) first. Under the hypothesis that Brazilian industry is close to a steady state in the mid eighties, the beginning-of-year foreign equipment stock in 1986 is set equal to $\underline{K}^{mach,*}_{86,i} = \sum_{s=86}^{88}(Acq^{mach,*}_{s,i}/Acq^{mach}_{s,i})/3 \cdot \underline{K}^{mach}_{86,i}$, where $Acq^{mach,*}_{t,i}$ and $\underline{K}^{mach,*}_{t,i}$ denote foreign equipment acquisitions and stocks, respectively. If a firm is recorded born before 1986 but appears in *PIA* only after 1986, the average share of foreign equipment acquisitions in the first two years of observations is used (instead of the three-year mean, as above). Turn to group (b) which contains new firms that enter *PIA* in or after 1986. If these firms are greenfield creations, their initial foreign equipment stock in 1986 is set to zero. If these firms have a legal or economic predecessor in *PIA*, the share of foreign equipment in the predecessor's total equipment stock in the year of succession is transferred to the successor as the adequate share of foreign equipment. If the firm is no greenfield creation but the predecessor is not observed in *PIA* in any previous year, the method of group (a) is applied.

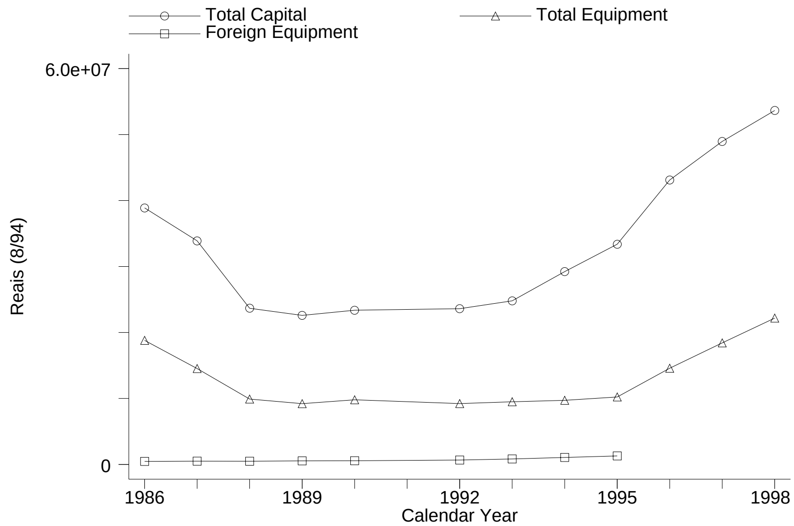
Step 5: The foreign equipment stock in all subsequent years, following the first year of observation of a firm, are inferred from the relationship $\overline{K}_{t,i}^{mach,*} = Acq_{t,i}^{mach,*} + \underline{K}_{t,i}^{mach,*}(1 - \hat{\sigma}_{t,i}^{mach} - \hat{\delta}_{t,i}^{mach})$. Under the assumption that a firm is equally likely to retire a domestic machine as it is to retire a foreign machine, the retirement of foreign equipment is approximated by $\hat{\sigma}_{t,i}^{mach}\underline{K}_{t,i}^{mach,*}$, where $\hat{\sigma}_{t,i}^{mach}$ is computed as $\hat{\sigma}_{t,i}^{mach} = Ret_{t,i}^{mach}/\underline{K}_{t,i}^{mach}$ ($Ret_{t,i}^{mach}$ denotes equipment retirements). Similarly, the assumption that foreign equipment depreciates at the same rate as domestic equipment is made and $\hat{\delta}_{t,i}^{mach}$ is calculated as in step 3. Finally, the problem to bridge the missing year 1991 occurs again. Applying similar arguments as in step 2, one can calculate $\hat{\phi}_{92,i}^{*} = \sum_{s=86}^{91}(\underline{K}_{s,i}^{mach,*}/\underline{K}_{s,i}^{mach})/6$ or an accordingly adjusted factor if years are missing (see step 2). Then, $\underline{K}_{92,i}^{mach,*} = \hat{\phi}_{92,i}^{*}\underline{K}_{92,i}^{mach}$. The remaining end-of-year stocks from 1992 until 1995 is inferred applying $\overline{K}_{t,i}^{mach,*} = Acq_{t,i}^{mach,*} + \underline{K}_{t,i}^{mach,*}(1 - \hat{\sigma}_{t,i}^{mach} - \hat{\delta}_{t,i}^{mach})$ again.

Wherever possible, missing values in *PIA*'s capital stock figures are imputed as $\overline{K}_{t,i} = (1 - \hat{\delta}_{t,i})\underline{K}_{t,i} + I_{t,i}$, using an estimate of the depreciation rate as in step 1 for total capital stock figures, and as in step 3 for the equipment stock. Throughout the construction of series for types of equipment, all components of the equipment stock are restricted to sum to the total.

A.5.11 Domestic equipment and its components

The domestic equipment stock can be split into further components until 1995. Vehicles, computers, and other capital goods are separately reported in *PIA velha*. According series are obtained with a procedure analogous to *Step 4* and *Step 5* in the preceding subsection. Contrary to the procedure for total assets and machinery, I do not apply the correction factor from section A.5.7 to vehicles, computers, and other capital goods. Similar to buildings, vehicles and computers behave differently than total assets and machinery before 1990 (1986-90). In the case of the computer stock, for instance, the computed correction factor from section A.5.7 would be 5.04. However, I use 3.5—the implied factor from accounting principles (table A.13)—since otherwise $\delta > 1$ for computers.

Figure A.5 shows the firm-average capital stock, equipment stock and foreign equipment stock as they result from the above efforts. Especially after the *Plano Real* stabilizes the economy in 1995, investment in the capital stock takes off. Foreign equipment is steadily accumulated from the late



Data: Unbalanced panel of all firms in *PIA* 1986-1998. Figures are unweighted sums.

Figure A.5: Firm-average capital stock, equipment and foreign equipment

1980s on.

A.5.12 Remarks on deflating liabilities

The correct valuation of liabilities in *PIA* remains an open issue. As discussed for the capital stock series, I play investment flows and depreciation rates against the stock series until I reach a mutually consistent series under a given set of reasonable criteria. There is no such choice for liability valuation since flows are not reported in *PIA* (and not recorded in a balance sheet in general). In addition, asset revaluations affect equity and thus the value of total liabilities. I therefore assess liability variables mainly through internal ratios such as the debt share in total liabilities, or the share of foreign short-term debt in total short-term liabilities and the like. Ratios such as liabilities per output would already pose a valuation problem that remains to be resolved. Some of these ratios can, surprisingly, exceed unity. The ratio of credit per total liabilities, for instance, can become larger than one since Brazilian accounting principles allow firms to show negative equity in their balance sheet temporarily.[13]

[13] Arguably, end-of-year values of economy-wide or industry-wide price indices could be applied to deflate the sum of credit, `crtot`. Since revaluations of assets, such as the optional revaluation programme in 1991, only affect the value of equity, the value of the sum of all credits would not be altered by this. Candidate economy-wide price indices to deflate the sum of credit (`crtot`) are *INPC* or *IGP-DI*. Just as in the case of flow variables, however, the use of a less general deflator may be more appropriate in the context of a firm's decision making process. For the firm, its decision to raise capital may depend on the relative prices of factors, and the relative sales price for final products. Therefore, another adequate deflator choice may be the use of industry-specific rather than economy-wide price indices. In particular, the use of the *IPA-OG* and *IPA-DI* series for deflating outputs and intermediate goods

A.6 Complementary Data

Mainly three additional data sources are exploited to complement information in *PIA*.

A.6.1 Market penetration series

Ramos and Zonenschain (2000b) present penetration series at the level of *nível 80*, comparable to the *SIC* three digit level. This sector grid is easily transformed to the one used in *PIA* (*nível 100*). The authors use monthly import and export data from the Brazilian national accounts in the period 1990-1998, and for 1980 and 1985. Among alternative sources of penetration rates, Ramos and Zonenschain's (2000b) series is most compatible with sector definitions in *PIA*. In addition, appealing data sources are available to the authors.

Foreign market penetration is measured as the share of imports in domestic absorption for a given sector. Call sector i's gross domestic output Y_i, and exports and imports EX_i and IM_i, respectively. Domestic absorption is $C_i + I_i + G_i \equiv A_i$ in standard notation for private consumption, investment and government consumption. A measure for market penetration can then be defined as

$$\frac{IM_i}{A_i} \equiv \frac{1}{\frac{Y_i-(EX_i-IM_i)}{IM_i}} = \frac{1}{1+\frac{1-\frac{EX_i}{Y_i}}{\frac{IM_i}{Y_i}}}. \tag{A.12}$$

For the decision of a firm that sells to the domestic market, this measure reflects the relevant competition more closely than the ratio of imports over output, say. Domestic firms find the absorption market (corresponding to A_i) the relevant environment in which they compete. Ramos and Zonenschain (2000b) provide the variables $\frac{EX_i}{Y_i}$ and $\frac{IM_i}{Y_i}$. Using the second equality in (A.12), the import penetration and export share series of Ramos and Zonenschain (2000b) are converted to this measure of foreign market penetration. Ramos and Zonenschain's (2000b) data points are 1980, 1985, and every year between 1990 and 1998. Import and export shares between 1986 and 1989 are missing. Since the economy remains fairly closed until 1990, it seems safe to infer missing years through linear interpolation.

A.6.2 Tariffs

Kume et al. (2000) report sector-specific tariff levels. They weigh product-specific tariffs with the value added in each narrowly defined product group and arrive at sector-specific tariff levels. Their sector classification (*nível 80*) is close to the one used in *PIA* (*nível 100*). The annual nominal tariffs used are simple annual means of the original monthly series. To compute input-side

inputs, suggest the use of the industry wide prices indices within the *IPA-OG* or *IGP-DI* series, too, to deflate the sum of credit, `crtot`.

tariffs for capital goods, the annual national capital formation vector is used as weighting scheme. For intermediate inputs, the national input-output matrix is used to construct annual and sector-specific weights. Both the national capital formation vector and the input-output matrix are obtained from *IBGE*'s national accounts.

A.6.3 Foreign direct investment

The Brazilian central bank publishes foreign direct investment (FDI) figures. On that basis, a series of FDI flows to Brazilian manufacturing sectors over the period 1986-1998 is constructed. Both the methodology and the sector definitions in the central bank data change in 1995. These changes cast doubt on the quality and consistency of FDI data up to 1995. However, a series of rough estimates can serve as a control in according regressions.

The cumulated FDI figures ('FDI stocks') for December 1995 are arguably most precise. The central bank conducted a survey among Brazilian firms in 1995. It seems that reported FDI stocks before 1995 do not match those of 1995. No FDI flows are available for 1995. However, under the assumption that aggregate FDI flows in 1995 equal more or less the average of the years 1993, 1994 and 1996, an approximate FDI flow of USD 1.71 Mio is inferred for 1995. With this estimate at hand, I compare the central bank's FDI series before and after 1995. The estimates of FDI stocks prior to 1995 turn out to be too high by a factor of 1.33. Therefore, all reported FDI stocks prior to 1995 are reduced by 1/1.33 and the implied flows are inferred accordingly. To make sector definitions consistent across the periods 1986-1995 and 1995-1998, and to make them compatible with *PIA*, sectors are aggregated to *nível 50* where possible.

A.6.4 Price indices of major trading partners

Price indices are obtained for Brazil's 25 major trading partners on the imports side in 1995. Their imports reached 89.8 percent of Brazil's total imports in 1995 and 91.2 percent in 1998. The top three countries alone supply 45.1 percent of Brazil's total imports in 1995 (USA 23.9 percent, Argentina 10.9, Germany 10.3). Wholesale price indices are used for Argentina, Chile, Italy, Japan, Mexico, Singapore, Taiwan, Uruguay, and Venezuela; producer prices for Belgium, Canada, France, Germany, Korea, Netherlands, Spain, Sweden, Switzerland, UK, and the US; and consumer prices for China, Hong Kong, Panama, Paraguay, and Saudi Arabia. The relative import share of these countries in 1995 is used as a fixed weight for average price indices in all years from 1986 through 1998.

A.7 Sectors of Industry

Firms in *PIA velha* are classified into sectors at the so-called *nível 100* (level 100). The definition of sectors of industry according to *nível 100* corresponds roughly to the three-digit *SIC* level in the US. *Nível 100* comes close to the sectoral definitions in the Brazilian national accounting system. However, the actual accounting system uses a classification system called *nível 80* which aggregates several manufacturing sectors in a slightly different way. Both *nível 100* and *nível 80* use a number system with four digits. The first two digits are identical in both systems (usually called *atividade 80*, *atividade 100*, or *nível 50*) and provide the simplest manner to move from *nível 100* to *nível 80*, and vice versa. However, it is possible to derive a finer mapping between sector definitions at *nível 80* and *nível 100*. Sectors 801 and 802, for instance, can be separated and correspond one-to-one to 810 and 820, respectively.

A.7.1 Compatibility between *Nível 100* and *CNAE*

Firms in *PIA nova* are classified according to a new system called *CNAE* (*Classificação Nacional de Atividades Empresariais*) which comes closer to international classifications. The following list shows how *CNAE* is transformed back to *nível 100* according to an internal recommendation at *IBGE*.

Nív.100	*CNAE*
210	1310, 1321, 1322, 1323, 1324, 1325, 1329
220	1410, 1421, 1429
310	1110, 1120
320	1000
410	2620
420	2630
430	2611, 2612, 2619
440	2641, 2642, 2649, 2691, 2692, 2699
510	2711, 2712, 2721, 2722, 2729
610	2741, 2742, 2749, 2752, 2832
710	2751, 2831
720	2731, 2739, 2811, 2812, 2833, 2834, 2839, 2841, 2842, 2843, 2891 2892, 2893, 2899
810	2813, 2821, 2822, 2911, 2912, 2913, 2914, 2915, 2921, 2922, 2923 2924, 2925, 2929, 2931, 2940, 2951, 2952, 2961, 2962, 2963, 2964 2965, 2969, 2971, 2972
820	2932, 2953, 2954
1010	3111, 3112, 3113, 3121, 3122
1020	3130, 3141, 3151, 3152, 3191
1030	2981, 2989, 3011, 3199
1110	3012, 3021, 3022, 3192, 3210, 3221, 3222, 3330
1120	3230
1210	3410, 3420, 3431, 3432, 3439

Nív.100	*CNAE*
1310	3142, 3160, 3441, 3442, 3443, 3444, 3449, 3450
1320	3511, 3512
1330	3521, 3522, 3523
1340	3531, 3532, 3591, 3592, 3599
1410	2010, 2021, 2022, 2023, 2029
1420	3611, 3612, 3613, 3614
1510	2110
1520	2121, 2122, 2131, 2132, 2141, 2142, 2149
1530	2211, 2212, 2213, 2214, 2219, 2221, 2222, 2229, 2231, 2232 2233, 2234
1610	2511, 2512, 2519
1710	2411, 2414, 2419, 2429
1720	2340
1810	2320
1820	2421, 2422
1830	2431, 2432, 2433, 2441, 2442
1910	2412, 2413
1920	2461, 2462, 2463, 2469, 2472, 2481, 2482, 2483, 2491, 2492, 2493 2494, 2496, 2499, 2310, 2330
2010	2451, 2452, 2453, 2454
2020	2471, 2473
2110	2521
2120	2522, 2529
2210	1711, 1719, 1721, 1722, 1731, 1732
2220	1723, 1733
2230	1724, 1741, 1749, 1750, 1761, 1762, 1763, 1764, 1769, 1771 1772, 1779
2310	1811, 1812, 1813, 1821, 1822
2410	1910, 1921, 1929
2420	1931, 1932, 1933, 1939
2510	1571, 1572
2610	1551
2620	1552
2630	1521, 1522, 1523, 1585
2640	1553, 1554, 1555, 1559, 1583
2650	1600
2710	1511, 1513
2720	1512
2810	1541, 1542
2910	1561, 1562
3010	1531
3020	1532, 1533
3110	1556
3120	1422, 1514, 1543, 1581, 1582, 1584, 1586, 1589
3130	1591, 1592, 1593, 1594, 1595
3210	2495, 3310, 3320, 3340, 3350, 3691, 3692, 3693, 3694, 3695, 3696 3697, 3699, 3710, 3720

A.7.2 Compatibility between *Nível 100*, *IPA-DI* and *IPA-OG*

The list below shows how the sectoral definition of *nível 100* are made compatible with the respective classifications in the price index series *IPA-DI* and *IPA-OG*. The list is joint work with Adriana Schor at Universidade de São Paulo. A list of the *IPA-DI* indices is given in subsection A.7.3 below.

Nív.100	*50*	Portuguese Description of Sector	*IPA-DI*	*IPA-OG*
210	2	Extração de minerais metálicos	mpr	28
220	2	Extração de minerais nao-metálicos	mpr	28
310	3	Extração de petróleo e gas natural	mpr	28
320	3	Extração de carvão mineral	mpr	28
410	4	Cimento e clínquer	constr	30
420	4	Peças e estruturas de concreto	constr	30
430	4	Vidro e artigos de vidro	mpr	30
440	4	Outros minerais não-metálicos	mpr	30
510	5	Siderurgia	mpr	32
610	6	Metalurgia dos não-ferrosos	mpr	33
710	7	Fundidos e forjados de aço	mpr	32
720	7	Outros produtos metalúrgicos	mpr	31
810	8	Máquinas, equipamentos e instalações	maq	36
820	8	Tratores e máquinas rodoviárias	maq	35
1010	10	Equipamentos para energia elétrica	maq	40
1020	10	Condutores e outros materiais elétricos	mpr	41
1030	10	Aparelhos e equipamentos elétricos	bcd-ud	39
1110	11	Material para aparelhos eletrônicos	mpr	38
1120	11	TV, radio, e equipamentos de som	bcd-ud	41
1210	12	Automóveis utilitários	veiculos	43
1310	13	Motores e peças para veículos	compveic	41
1320	13	Indústria naval	trans	44
1330	13	Indústria ferroviária	trans	44
1340	13	Fabricação de outros veículos	trans	43
1410	14	Indústria da madeira	mpr	45
1420	14	Indústria do mobiliário	bcd-ud	46
1510	15	Celulose e pasta mecânica	mpr	50
1520	15	Papel, papelão e artefatos de papel	mpr	50
1610	16	Indústria da borracha	mpr	51
1710	17	Elementos químicos não petroquímicos	mpr	58
1720	17	Destilação de álcool	mpr	54
1810	18	Refino de petróleo	mpr	54
1820	18	Petroquímica	mpr	58
1830	18	Resinas, fibras e elastomeros	mpr	56
1910	19	Adubos e fertilizantes	mpr	57
1920	19	Produtos químicos diversos	mpr	53
2010	20	Indústria farmac eutica	bcnd	81[a]
2020	20	Indústria de perfumaria, sabões e velas	bcnd	82[a]
2110	21	Laminados plásticos	mpr	83[a]
2120	21	Artigos de material plástico	bcnd	83
2210	22	Beneficiamento de fibras naturais	mpr	60
2220	22	Fiação de fibras artificiais	mpr	61

Nív.100	*50*	Portuguese Description of Sector	*IPA-DI*	*IPA-OG*
2230	22	Outras indústrias téxteis	mpr	65[a]
2310	23	Artigos do vestuario e acessórios	bcnd	63
2410	24	Indústria de couros e peles	mpr	52
2420	24	Calçados	bcnd	64
2510	25	Indústria do café	bcnd-alim	75[a]
2610	26	Beneficiamento do arroz	bcnd-alim	76[a]
2620	26	Moagem de trigo	mpr	72
2630	26	Conservação de frutas e legumes	bcnd-alim	76
2640	26	Outros produtos vegetais	bcnd-alim	76
2650	26	Indústria do fumo	bcnd	69
2710	27	Preparação de carnes	bcnd-alim	78
2720	27	Preparação de aves	bcnd-alim	78
2810	28	Preparação do leite e laticínios	bcnd-alim	79
2910	29	Indústria do açucar	bcnd-alim	73
3010	30	Óleos vegetais em bruto	mpr	74
3020	30	Refino de óleos vegetais	bcnd-alim	74
3110	31	Alimentos para animais	mpr	80
3120	31	Outras indústrias alimentícias	bcnd-alim	80
3130	31	Indústria de bebidas	bcnd-alim	66
3210	32	Outras indústrias	ipadi	29

[a]The price index series *IPA-OG* 65, 75, 81, 82, and 83 begin in March 1986, and *IPA-OG* 76 in January 1970. Their earlier years are replaced with according aggregate indices, rebased to the connecting year: 65 with 59, 75 and 76 with 71, and 81 through 83 with 29.

A.7.3 Categories of *IPA-DI* price index series

The abbreviations for *IPA-DI* price indices are explained in the table below. As the table shows, several aggregate categories of indices are not used.

Category	*IPA-DI* series (Portuguese description)
ipadi	Total - Média Geral
.	Bens de Consumo - Total
bcd	Bens de Consumo Duráveis - Total
.	Bens de Consumo Duráveis - Outros
bcd-ud	Bens de Consumo Duráveis - Utilidades Domésticas
bcnd	Bens de Consumo Não Duráveis - Total
bcnd-alim	Bens de Consumo Não Duráveis - Gêneros Alimentícios
.	Bens de Consumo Não Duráveis - Outros
.	Bens de Produção - Total
compveic	Bens de Produção - Componentes para Veículos[a]
.	Bens de Produção - Máquinas, Veículos e Equipamentos, Total
maq	Bens de Produção - Máquinas e Equipamentos
veic	Bens de Produção - Veículos Pesados para Transporte
constr	Bens de Produção - Materiais de Construção
mpr	Bens de Produção - Matérias Primas, Total
.	Bens de Produção - Matérias Primas Brutas
.	Bens de Produção - Matérias Primas Semi-Elaboradas
.	Bens de Produção - Outros

Category	*IPA-DI* series (Portuguese description)
veiculos	Unweighted mean of *bcd* and *veic*
	Bens de Consumo Duráveis - Total
	Bens de Produção - Veículos Pesados para Transporte
trans	Unweighted mean of *compveic* and *veic*
	Bens de Produção - Componentes para Veículos[a]
	Bens de Produção - Veículos Pesados para Transporte

[a]Only since 1986.

A.7.4 English descriptions of sectors at *Nível 80*

A list of *IBGE*'s English descriptions of sectors at *nível 80* is given below.

Nív.80	*Nív.50*	English Description of Sector
201	2	Iron ore mining
202	2	Mining of other metals
301	3	Oil and gas production
302	3	Coal and other mining
401	4	Non-metallic mineral products
501	5	Basic metallic products
502	5	Rolled steel
601	6	Non-ferrous metallic products
701	7	Other metallic products
801	8	Manufacturing and maintenance of machinery and equipment
802	8	Tractors and embankment machinery
1001	10	Electrical equipment
1101	11	Electronic equipment
1201	12	Automobiles, trucks, and buses
1301	13	Other vehicles and parts
1401	14	Timber and furniture
1501	15	Paper, pulp, and cardboard
1601	16	Rubber products
1701	17	Non-petrochemical chemical elements
1702	17	Alcohol
1801	18	Motor gasoline
1802	18	Fuel oil
1803	18	Other refinery products
1804	18	Basic petrochemical products
1805	18	Resins and fibers
1806	18	Alcoholic fuel
1901	19	Chemical fertilizers
1902	19	Paints, varnishes, and lacquers
1903	19	Other chemical products
2001	20	Pharmaceutical products and perfumes
2101	21	Plastics
2201	22	Natural textile fibers
2202	22	Natural textiles
2203	22	Artificial textile fibers
2204	22	Artificial textiles

Nív.80	*Nív.50*	English Description of Sector
2205	22	Other textile products
2301	23	Apparel
2401	24	Leather products and footwear
2501	25	Coffee products
2601	26	Processed rice
2602	26	Wheat flour
2603	26	Other processed edible products
2701	27	Meat
2702	27	Poultry
2801	28	Processed milk
2802	28	Other dairy products
2901	29	Sugar
3001	30	Raw vegetable oil
3002	30	Processed vegetable oil
3101	31	Animal food and other food products
3102	31	Beverages
3201	32	Miscellaneous

A.8 Geographic Regions of Brazil

Firms are grouped by region. *PIA* follows the principle to list a firm in the region where the legal headquarters of the firm is located. This need not be the region where the firm creates most value. The following list gives an overview of the regions (variable `region`) and their codes, and the number of observations for each region and state (`uf`, *Unidade Federal*).

State	`uf`	Name	Valid Obs. [a]	Percent
`region` 1: North (*Norte*)			2,166 [b]	2.75 [b]
RO	11	*Rondônia*	253	.32
AC	12	*Acre*	63	.08
AM	13	*Amazonas*	839	1.47
RR	14	*Roraima*	0	.00
PA	15	*Pará*	956	1.22
AP	16	*Amapá*	35	.04
TO	17	*Tocantins*	17	.02
`region` 2: North-East (*Nordeste*)			7,483 [b]	9.51 [b]
MA	21	*Maranhão*	455	.58
PI	22	*Piauí*	190	.24
CE	23	*Ceará*	1,331	1.69
RN	24	*Rio Grande do Norte*	561	.71
PB	25	*Paraíba*	523	.67
PE	26	*Pernambuco*	1,855	2.36
AL	27	*Alagoas*	460	.59
SE	28	*Sergipe*	321	.41
BA	29	*Bahia*	1,762	2.24

State	`uf`	Name	Valid Obs.[a]	Percent
`region` 3: South-East (*Sudeste*)			50,117[b]	63.67[b]
MG	31	*Minas Gerais*	6,042	7.69
ES	32	*Espírito Santo*	956	1.22
RJ	33	*Rio de Janeiro*	6,753	8.59
SP	35	*São Paulo*	36,259	46.15
`region` 4: South (*Sul*)			17,084[b]	21.70[b]
PR	41	*Paraná*	4,821	6.14
SC	42	*Santa Catarina*	4,316	5.49
RS	43	*Rio Grande do Sul*	7,946	10.11
`region` 5: Center-West (*Centro-Oeste*)			1,863[b]	2.37[b]
MS	50	*Mato Grosso do Sul*	373	.47
MT	51	*Mato Grosso*	377	.48
GO	52	*Goías*	840	1.07
DF	53	*Distrito Federal*	264	.34
Subtotal			78,571	100.00
Unclassified			394	
Total			78,965	

[a] Observations with `catlife` equal to 9.3, 9.35, or 9.99 removed

[b] Observations for `region` are independent of `uf` (*Subtotal* of regions: 78,713).

A.9 Detailed Categories of a Firm's 'Economic Curriculum'

This section presents fine rosters to classify firms according to their 'economic curriculum.' The first subsection A.9.1 is dedicated to categories of entry, whereas the second subsection A.9.2 deals with both the life (possible periods of suspended production) and the type of exit of a firm. The rosters are presented along with the algorithms to classify the firms in *PIA*. The categories are grouped according to four-digit arabic numbers, and more detailed instructions about applicable algorithms are given either with the definition of the category or in brackets. The algorithms mainly draw on the variables `state` and `change` and on whether a firm reports positive sales in a given year or not.

Useful additional pieces of information are the effective founding year of a firm (`effborn`, see section A.2.6 and upper part of table A.9) and whether a firm is continuously present in *PIA* or not. For the latter, an auxiliary variable called `contgrp` is created. The variable `contgrp` takes four possible values

1: Continuous presence in all sample years

2: Continuous presence until apparently early exit from sample
[missing years at end of *PIA* only]

3: Continuous presence after apparently late entry into sample
[missing years at beginning of *PIA* only]

4: Interrupted presence [missing years at some other point]

Here, presence in a year means strictly positive sales in that year.

A.9.1 Detailed categories of entry

Categories marked with an asterisk draw on information flowing from the 'family tree' of firms (see section A.2.4). Conditions for higher-order groups apply to all lower-order groups.

1: Old firm that appears in *PIA* in 1986 or later
[`effborn` < 1986]

(2): New and 'well born' firm during sample period
[`effborn`=year of first appearance]

- *2.1: Baby firm ('Greenfield creation')
[firm does not satisfy criteria for categories 2.2-2.5 of `catentr`]
- *2.2: Creation as Legal Successor of existing firm (mere change of *tax number* or absorption by other firm)
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), and firm does not satisfy criteria for any of the following categories of `catentr`, 2.3-2.5]
- *2.3: Creation through Merger of existing firms
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), referencing 'parent' records `change`=1]
- *2.4: Creation through complete Split-Up of existing firm
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), referencing 'parent' records `change`=4 or 5]
- *2.5: Creation as Spin-Off of existing firm
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), referencing 'parent' records `change`=6]

3: Apparently new born firm in *PIA* (`state`=2 in *PIA*), but not reported in register
[`effborn` missing, but `state`=2 in first year of appearance]

(4): New born firm, but lag before appearance in *PIA* (lag of no more than 3 years)

- (4.1): Lag of 1 year between registration in tax or *IBGE*'s register and first appearance in *PIA*
[`effborn` 1 year before first appearance]
  - *4.11: Baby firm
[firm does not satisfy criteria for any of the following categories of `catentr`, 4.12-4.15]
  - *4.12: Creation as Legal Successor of existing firm
[firm born after year of being referenced by 'parent' firm (`effborn` >=year of referencing), and firm does not satisfy criteria for any of the following categories of `catentr`, 4.13-4.15]
  - *4.13: Creation through Merger of existing firms
[firm born after year of being referenced by 'parent' firm (`effborn` >=year of referencing), referencing 'parent' records `change`=1]

*4.14: Creation through complete Split-Up of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=4 or 5]

*4.15: Creation as Spin-Off of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=6]

(4.2): Lag of 2 years between registration in tax or *IBGE*’s register and first appearance in *PIA*
[`effborn` 2 years before first appearance]

*4.21: Baby firm
[firm does not satisfy criteria for any of the following categories of `catentr`, 4.22-4.25]

*4.22: Creation as Legal Successor of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), and firm does not satisfy criteria for any of the following categories of `catentr`, 4.23-4.25]

*4.23: Creation through Merger of existing firms
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=1]

*4.24: Creation through complete Split-Up of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=4 or 5]

*4.25: Creation as Spin-Off of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=6]

(4.3): Lag of 3 years between registration in tax or *IBGE*’s register and first appearance in *PIA*
[`effborn` 3 years before first appearance]

*4.31: Baby firm
[firm does not satisfy criteria for any of the following categories of `catentr`, 4.32-4.35]

*4.32: Creation as Legal Successor of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), and firm does not satisfy criteria for any of the following categories of `catentr`, 4.33-4.35]

*4.33: Creation through Merger of existing firms
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=1]

*4.34: Creation through complete Split-Up of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=4 or 5]

*4.35: Creation as Spin-Off of existing firm
[firm born after year of being referenced by ‘parent’ firm (`effborn` >=year of referencing), referencing ‘parent’ records `change`=6]

7: Late comer: Old firm that only appears in *PIA* later than 1986 (foundation strictly more than three years earlier)
[`effborn` more than 3 years before first appearance]

(8): Out of the blue: Firm without age (no entry in tax or *IBGE*’s register) or birth
[`effborn` empty and `state`$\neq$2]

*8.1: Truly out of the blue
[firm does not satisfy criteria for categories 8.2-8.5 of `catentr`]

*8.2: 'Family tree' allows classification as Legal Successor of existing firm
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), and firm does not satisfy criteria for any of the following categories of `catentr`, 8.3-8.5]

*8.3: 'Family tree' allows classification as Merger of existing firms
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), referencing 'parent' records `change`=1]

*8.4: 'Family tree' allows classification as Successor from Split-Up
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), referencing 'parent' records `change`=4 or 5]

*8.5: 'Family tree' allows classification as Successor from Spin-Off
[firm born after year of being referenced by 'parent' firm (`effborn`>=year of referencing), referencing 'parent' records `change`=6]

(9): Differently behaved firms

9.1: Firm enters like young (installation process), but is old according to tax or *IBGE*'s register
[`effborn` earlier than first appearance, but `state`=2 at first appearance]

9.2: Birth according to tax or *IBGE*'s register later than first appearance in *PIA*
[`effborn` later than first appearance]

9.3: Installation process observed after first appearance in *PIA*
[`state`=2 in a year strictly later than year of first appearance]

A.9.2 Detailed categories of exit and suspended production

Categories marked with an asterisk draw on information flowing from the 'family tree' of firms (see section A.2.4). Conditions for higher-order groups apply to all lower-order groups. Categories in curly brackets are never assigned by only listed here to clarify the classification system.

0: No exit, no period of suspended production, or missing sales observed after first appearance in sample
[both `state`<=2 and strictly positive sales in every year after first appearance (or `state`=2, 5 or 8 and `contgrp`=3)]

(1): Complete absorption by other firm
[`state`=4 or 6 (or `state`=8 and `change`=1, 2, 4, or 7; or `state`=1, `change`=10, and year=last year of appearance) and *tax number link* is set]

1.1: Change of legal status (inferred or from data)
[firm not `catlife`=1.2, 1.3, or 1.4, and successor born in year of referencing]

1.2: Merger [`change`=1]

1.3: Acquisition by existing firm
[firm not `catlife`=1.4 and successor born before referenced]

1.4: Delayed acquisition after complete suspension or exit
[at least one year with `state`=5, 6, or 8 and no sales after suspension period or exit, then acquisition by other firm]

2: Exit
[`state`=4 or 6 (or `state`=5 and year=last year of appearance; or `state`=8, `change` not set, no sales, and year is last year of appearance) and no *tax number link* set]

(3): Temporarily suspended production during sample period
[state=3 or 5 (or state=8, no sales, and change=8; or state=8, no sales, no successor, and change empty)]

3.0: No absorption or exit in any later period
[firm satisfies none of criteria for catlife 3.11-3.2]

3.11: Change of legal status in distant later period (at least 1 year of observed operation inbetween)
[firm satisfies criteria of catlife 1.1 otherwise]

3.12: Merger in distant later period (at least 1 year of observed operation inbetween)
[firm satisfies criteria of catlife 1.2 otherwise]

3.13: Acquisition by existing firm in distant later period period (at least 1 year of observed operation inbetween)
[firm satisfies criteria of catlife 1.3 otherwise]

3.14: Delayed acquisition in distant later period period after complete suspension (at least 1 year of observed operation inbetween)
[firm satisfies criteria of catlife 1.4 otherwise]

3.2: Exit in distant later period (at least 1 year of observed operation inbetween)
[firm satisfies criteria of catlife 2 otherwise]

(5): Missing data

5.0: Missing years
[Missing year(s) but firm satisfies none of criteria for catlife 5.1 through 5.3]

(5.1): Missing years before complete absorption by other firm (effective exit year adjusted accordingly)
[Missing year(s) before absorption. state=4 or 6 (or state=8 and change=1, 2, 4, or 7; or state=1, change=10, and year=last year of appearance) and *tax number link* is set]

5.11: Missing years immediately before change of legal status
[firm satisfies criteria of catlife 1.1 otherwise]

5.12: Missing years immediately before merger
[firm satisfies criteria of catlife 1.2 otherwise]

5.13: Missing years immediately before acquisition by existing firm
[firm satisfies criteria of catlife 1.3 or 5.5 otherwise]

5.14: Missing years immediately before ailing to delayed acquisition starts
[firm satisfies criteria of catlife 1.4 otherwise]

5.2: Missing years immediately before exit
[Missing year(s) before exit. state=4 or 6 (or state=5 and year=last year of appearance; or state=8, change not set, no sales, and year=last year of appearance) and no *tax number link* set]

(5.3): Missing years in neighboring year to period of suspended production (years imputed with state=9)
[Missing year(s) during period of suspended production and firm does not simultaneously satisfy criteria for catlife 5.1. In addition, state=3 or 5 (or state=8, no sales, and change=8; or state=8, no sales, no successor, and change empty)]

5.30: and no absorption or exit in any later period
[firm satisfies none of criteria for catlife 5.311-5.32]

5.311: and change of legal status in distant later period
[firm satisfies criteria of catlife 3.11 otherwise]

5.312: and merger in distant later period
[firm satisfies criteria of `catlife` 3.12 otherwise]

5.313: and acquisition in distant later period
[firm satisfies criteria of `catlife` 3.13 otherwise]

5.314: and delayed acquisition in distant later period
[firm satisfies criteria of `catlife` 3.14 otherwise]

5.32: and exit in distant later period
[firm satisfies criteria of `catlife` 3.2 otherwise]

5.5: Missing age of acquiring firm does not permit distinction of 1.1 and 1.3
[`state`=4 or 6 (or `state`=8, no sales, and `change`=1, 2, 4, or 7; or `state`=1, `change`=10, and year=last year of appearance) and *tax number link* is set; in addition, firm not `catlife`=1.2 or 1.4 and `effborn` not known for referenced successor firm]

(8): Not elsewhere categorized

8.0: Missing sales in at least one period, next best category 3.0
[in at least one year `state`=1 but no sales and no successor, and in every year `change` empty or `change`=10]

8.1: Combinations of `change`=10 and successor firm indicate possible name `change`, next best category 1.1
[firm does not satisfy criteria of any other `catlife` category in 1-5 or 9, `change`=10, and *tax number link* set (`state` may take any value)]

8.2: Combinations of `state`=8 and `change`=10 and no successor firm make firm fall through previous roster
[firm does not satisfy criteria of any other `catlife` category in 1-5, 8.0, 8.1 or 9, `state`=8, `change`=10, and *tax number link* not set]

8.3: Combinations of `state`=8 and `change`=? or `state`=? and `change`=10 make firm fall through previous roster
[firm does not satisfy criteria of any other `catlife` category in 1-5, 8.0-8.2 or 9, `state`=8, or `change`=10, or both]

8.7: Firm being non-industrial in at least one period (`state`=7) makes it fall through previous roster
[firm does not satisfy criteria of any other `catlife` category in 1-5, 8.0-8.3 or 9, and `state`=7 in at least one year]

(9): Contradictory or Problematic Exiting or Standstill Behavior

9.1: Firm is marked extinct but lives on or reappears
[`state`=4 or 6 (or `state`=8 and `change`=1, 2, 4, or 7) in some year, but strictly positive sales recorded in a later year]

*9.15: Firm may be put back to better category due to cross-referencing

9.2: Firm is marked as in built-up phase but was working before
[`state`=2 in some year but strictly positive sales in an earlier year]

9.3: Effective year of exit is year of first appearance in *PIA* or no sales ever
[`effextyr`<=first year of appearance, or no strictly positive sales in any year]

*9.35: Firm may be put back to better category due to cross-referencing

9.99: Firm never found manufacturing in PIA
[firm does not satisfy criteria for `catlife`=9.3; and `state`>=5 in every year]

A.10 Economic Variables in *PIA*

Table A.14 documents the manner in which I construct consistent economic variables. The numbers in columns 3 through 5 indicate the 'id number' of the variables in the respective years of *PIA*. The 'id numbers' in columns 3 and 4 are precisely the numbers of the fields in the questionnaires of *PIA velha*. Due to the fact that two types of questionnaires exist in *PIA nova*, the id number in column 5 of table A.14 is only equal to the field in the questionnaire when the id number is not preceded by an 'x'. The according translation from 'x'-ed variables into the id numbers in the long questionnaire (*questionário completo*) are given below table A.14.

Some economic variables are inherently hard to deflate, such as liabilities. A simple way to use these variables but to avoid deflation problems is to express the liability structure through ratios. Similarly, social contributions and benefits may be hard to deflate, and it appears preferable to express their relation to total expenditures for personnel in ratios. Table A.15 summarizes possible definitions for such ratios that are consistent over time. It also includes the ratio of foreign intermediate goods purchases per total intermediate goods purchases. This variable is reported in *PIA nova* since 1996.

Table A.14: Economic Variables

Variable	Description	PIA 86-90	PIA 92-95	PIA 96-98[a]
`grssales`	Gross Sales of Final Goods	103	56	x15+16
`sales`	Net Sales of Final Goods	$109 \cdot \frac{103+105}{103+104+105}$	$62 \cdot \frac{56+58}{56+57+58}$	$\text{x14} \cdot \frac{\text{x15+16}}{\text{14+15+16}}$
`difstock`[b]	Change in Processed Goods Stocks	142	96	43-47+44-48
`resales`	Resales of Merchandise	104	57	15
`intmacq`[c]	Acquis. of Intermediate Goods	140	94	x26+58+63+71
`intmdif`[b]	Change in Interm. Goods Stocks	141	95	42-46
`labftot`	Labor (Total)	28	27	x01
`labftop`[d]	Labor Force: Top Management	24	24	x07
`labfwh`[d]	Labor Force: White-Collar	25	25	x05
`labfbl`	Labor Force: Blue-Collar	26	26	x03
`wagetot`	Salaries (Total)	33	32	x09
`wagetop`[d]	Salaries: Top Management	29	28	x12
`wagewh`[d]	Salaries: White-Collar	$30 + 32 \cdot \frac{30}{30+31}$	$29 + 31 \cdot \frac{29}{29+30}$	x11
`wagebl`	Salaries: Blue-Collar	$31 + 32 \cdot \frac{31}{30+31}$	$30 + 31 \cdot \frac{30}{29+30}$	x10
`astot`	Assets (Total)	11	11	.
`asliq`	Liquid Assets (in Total)	1	1	.

[a] Variables with a preceding 'x' indicate variables in *PIA nova* that have different names in questionnaires *quesionário completo* and *simplificado*. The x-variables correspond to the following variables in *questionário completo*: x01:=4, x03:=1, x05:=2, x07:=3, x09:=12, x10:=9, x11:=10, x12:=11, x14:=20, x15:=14, x26:=40.

[b] Initial stock less final stock.

[c] Includes electricity consumption and expenditure for equipment repair.

[d] Not strictly compatible between *PIA velha* (1986-95) and *PIA nova* (from 1996 on). Difference in classification of senior managers. See section A.3.1.

Table A.14: Economic Variables, continued

Variable	Description	PIA 86-90	PIA 92-95	PIA 96-98[a]
`aslr`	Long-run Assets (in Total)	6	6	.
`aspsum`	Permanent Assets (Sum; in Long-run A.)	7	7	.
`aspinv`	Perm A.: Holdings of Investments	8	8	.
`aspimmo`	Perm A.: Equipment & Real Estate	9	9	.
`aspmasum`	Perm A.: Machinery (in Eq.&R.Est.)	97	.	.
`aspdefer`	Perm A.: R&D & Fiscal Operations	10	10	.
`asrtimmo`	Rental of Equipment & Real Est.	132	86	x36
`aslsimmo`	Leasing of Equipment	133	87	x37
`deprec`	Asset Depreciation Cost	135	89	61
`fincost`[b]	Financial Costs	117	70	67+68-28
`acqtot`	Acquisitions of Assets (Total)	56	47	80+85+x53
`acqbl`	Acquisition of Ground & Premises	42+43	33+34	x55+x59+x63
`acqmasum`	Acquisitions of Machinery (Sum)	46	37	x56+x60+x64
`acqmadom`	Acquis. of Machinery: Domestic	47	38	.
`acqmause`	Acquis. of Machinery: Used	49	40	.
`acqmafor`	Acquis. of Machinery: Foreign	48	39	.
`acqveh`	Acquisitions of Vehicles	50	41	x57+x61+x65
`acqother`	Acquisitions of Other Assets	53+54+55	44+45+46	x58+x62+x66
`acqcomp`	Acquis. of Other Ass.: Computers	54	45	.

[a] Variables with a preceding 'x' indicate variables in *PIA nova* that have different names in questionnaires *quesionário completo* and *simplificado*. The x-variables correspond to the following variables in *questionário completo*: x36:=59, x37:=60, x53:=90, x55:=76, x56:=77, x57:=78, x58:=79, x59:=81, x60:=82, x61:=83, x62:=84, x63:=86, x64:=87, x65:=88, x66:=89.

[b] Includes costs and benefits from monetary correction.

Table A.14: Economic Variables, continued

Variable	Description	PIA 86-90	PIA 92-95	PIA 96-98[a]
`asltot`	Sales of Assets (Total)	72	55	x54
`aslbl`	Sales of Ground & Premises	65+66	48+49	x67
`aslmasum`	Sales of Machinery	67	50	x68
`aslveh`	Sales of Vehicles	68	51	x69
`aslother`	Sales of Other Assets	69+70+71	52+53+54	x70
`aslcomp`	Sales of Other Ass.: Computers	70	53	.
`balsum`[b]	Total Liabilities	23	23	.
`crtot`[c]	Credit (Total)	12+17	12+17	.
`crstsum`[d]	Short-Term Credit (Sum)	12	12	.
`crstdom`[d]	Short-Term Credit: Domestic	14	14	.
`crstfor`[d]	Short-Term Credit: Foreign	15	15	.
`crltsum`[d]	Long-Term Credit (Sum)	17	17	.
`crltdom`[d]	Long-Term Credit: Domestic	18	18	.
`crltfor`[d]	Long-Term Credit: Foreign	19	19	.
`profit`[e]	Profit before tax	126-127+124+125	80-81+77+78+79	74-75

[a] Variables with a preceding 'x' indicate variables in *PIA nova* that have different names in the questionnaires *quesionário completo* and *simplificado*. The x-variables correspond to the following variables in *questionário completo*: x54:=95, x67:=91, x68:=92, x69:=93, x70:=94.

[b] Since asset revaluations affect equity, this variable is extremely hard to value. It is therefore only used in ratios. See table A.15.

[c] Industry-wide prices indices within the *IPA-OG* or *IGP-DI* series or economy-wide price indices may arguably be adequate deflators.

[d] Reliable deflation methods remain to be developed. This variable is used in ratios only. See table A.15.

[e] The proposed figure is not strictly compatible before and after 1990. Social contributions under *lei 7689 de 15/12/1988* reduce the profits in addition to the tax payments from 1989 on. This fact is only accounted for after 1991. So, the years 1989 and 1990 are not strictly consistent with the other years. Also see section A.3.1 on this.

Table A.15: Ratios of Economic Variables

Variable	Description	PIA 86-90	PIA 92-95	PIA 96-98[a]
`crrat`	Ratio: Credit in Balance Sum	(12+17)/23	(12+17)/23	.
`crltrat`	Ratio: Long-Term Cr./Tot. Credit	17/(12+17)	17/(12+17)	.
`crforrat`	Ratio: Foreign Cr./Total Credit	(15+19)/(12+17)	(15+19)/(12+17)	.
`crstfrat`	Ratio: Foreign Short-Tm./STm Cr.	15/12	15/12	.
`crltfrat`	Ratio: Foreign Long-Term/LTm Cr.	19/17	19/17	.
`deprcrat`	Ratio: Depreciation/Total Cost	135/*Total Cost*[b]	89/*Total Cost*[c]	61/(x33-x41)
`finexrat`[d]	Ratio: Financial Cost/Total Cost	117/*Total Cost*[b]	70/*Total Cost*[c]	(67+68-28)/(x33-x41)
`sallcrat`	Ratio: Salaries/Total Labor Cost	$\frac{128+129}{128+129+130+131}$	$\frac{82+83}{82+83+84+85}$	x43/x42
`soclcrat`[e]	Ratio: Soc.Contrib./Tot.Lab.Cost	$\frac{130}{128+129+130+131}$	$\frac{84}{82+83+84+85}$	$\frac{x44+x45+x46+x47}{x42}$
`benlcrat`[f]	Ratio: Benefits/Total Labor Cost	$\frac{131}{128+129+130+131}$	$\frac{85}{82+83+84+85}$	x48/x42
`labtcrat`	Ratio: Labor Cost/Total Cost	$\frac{128+129+130+131}{Total\ Cost}$[g]	$\frac{82+83+84+85}{Total\ Cost}$[h]	x42/(x42+x33-x41)
`intfrrat`	Ratio: Foreign Intm./Tot. Intm.	.	.	51[i]

[a] Variables with a preceding 'x' indicate variables in *PIA nova* that have different names in questionnaires *quesionário completo* and *simplificado*. The x-variables correspond to the following variables in *questionário completo*: x33:=73, x41:=72, x42:=39, x43:=33, x44:=34, x45:=35, x46:=36, x47:=37, x48:=38.

[b] *Total Cost*: 119+132+133+134+135+138+139+140. 117 not included to avoid double count.

[c] *Total Cost*: 72+86+87+88+89+92+93+94. 70 not included to avoid double count.

[d] Includes costs and benefits from monetary correction.

[e] Social contributions include payments to the federal Brazilian social security system, to private pension funds, to health insurances and care providers.

[f] Benefits include: Transport, board, educational programs, day nurseries, and the like.

[g] *Total Cost*: 119+128+129+130+131+132+133+134+135+138+139+140. 117 not included to avoid double count.

[h] *Total Cost*: 72+82+83+84+85+86+87+88+89+92+93+94. 70 not included to avoid double count.

[i] Named *PERCEST*. Original figure is percentage.

Appendix B

Mathematical appendix to chapter 4

The present appendix provides the mathematical background to derivations in chapter 4.

B.1 Conjugate prior distributions

The framework of chapters 4 and 5 seems to be suited for an application to further conjugate prior distributions. Raiffa and Schlaifer (1961) provide an early overview of conjugate prior distributions. They discuss the beta-binomial, normal-normal, gamma-Poisson, and gamma-exponential pairs. Robert (1996, Ch. 3.2.3) lists the gamma-gamma, beta-Negative binomial, Dirichlet-multinomial, and gamma-normal pairs in addition.

Suppose signals $s_j \in \mathbb{R}^k$ are distributed with $f(s_j|\theta)$.

Definition B.1 *A family of probability distributions on a parameter space $\Theta \subseteq \mathbb{R}^k$ is said to be* conjugate *(or* closed under sampling*) if, for every prior distribution $\pi(\theta)$ in this family, the posterior distribution $\pi(\theta|s_1, ..., s_N)$ belongs to the same family.*

Conjugate prior distributions have many nice properties and are closely related to exponential families (see e.g. Brown (1986)). In particular,

Proposition B.1 *If $f(s_j|\theta)$ belongs to a natural exponential family, then, for any sample of signals $s_1, ..., s_N \overset{i.i.d.}{\sim} f(s_j|\theta)$, there exists a sufficient statistic of constant dimension*

$$\overline{s} = \frac{1}{N}\sum_{j=1}^{N} s_j \in \mathbb{R}^k$$

for all N.

Proof. See Brown (1986). ■

The following converse is also true.

Proposition B.2 *If a family of distributions $f(\cdot|\theta)$ is such that, for a sample size large enough, there exists a sufficient statistic of constant dimension, the family is exponential if the support of $f(\cdot|\theta)$ does not depend on θ.*

Proof. See Jeffreys (1939, §3.71). ■

The restriction on the support of $f(\cdot|\theta)$ is necessary for the converse to hold since the uniform and the Pareto distributions also satisfy this property.

Moreover, conditional independence is satisfied whenever a conjugate prior distribution is applied.

B.2 Properties of the gamma and Poisson distributions

The following statements summarize useful properties of the gamma distribution.

Fact B.1 *Mean and variance of θ are $\mathbb{E}[\theta\,|\alpha,\beta] = \alpha/\beta$ and $\mathbb{V}(\theta\,|\alpha,\beta) = \alpha/\beta^2$, respectively.*

Proof. The k^{th} raw moment of a gamma distributed random variable Z is

$$\mathbb{E}\left[Z^k\,|\alpha,\beta\right] = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}.$$

See DeGroot (1989, Ch. 5.9). This immediately establishes fact B.1. ■

Fact B.2 *For an arbitrary constant c, the expected value of $e^{-c\cdot\theta}$ is*

$$\mathbb{E}\left[e^{-c\cdot\theta}\,|\alpha,\beta\right] = \left(\frac{\beta}{\beta+c}\right)^{\alpha}.$$

Proof. Write the expected value of e^{-cZ} as

$$\begin{aligned}
\mathbb{E}\left[e^{-cZ}\,|\alpha,\beta\right] &= \int_0^\infty \frac{(\beta)^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-(\beta+c)z}\,\mathrm{d}z \\
&= \left(\frac{\beta}{\beta+c}\right)^{\alpha} \int_0^\infty \frac{(\beta+c)^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-(\beta+c)z}\,\mathrm{d}z = \left(\frac{\beta}{\beta+c}\right)^{\alpha},
\end{aligned}$$

where the last step follows from the property of any p.d.f that $\int f(z)\,\mathrm{d}z = 1$. ■

Fact B.3 *For an arbitrary constant c, the expected value of* $\theta \cdot e^{-c\cdot\theta}$ *is*

$$\mathbb{E}\left[\theta e^{-c\cdot\theta}\,|\alpha,\beta\right] = \left(\frac{\beta}{\beta+c}\right)^{\alpha}\frac{\alpha}{\beta+c}.$$

Proof. Write the expected value of Ze^{-cZ} as

$$\begin{aligned}\mathbb{E}\left[Ze^{-cZ}\,|\alpha,\beta\right] &= \left(\frac{\beta}{\beta+c}\right)^{\alpha}\int_0^{\infty} z\frac{(\beta+c)^{\alpha}}{\Gamma(\alpha)}z^{\alpha-1}e^{-(\beta+c)z}\,\mathrm{d}z \\ &= \left(\frac{\beta}{\beta+c}\right)^{\alpha}\frac{\alpha}{\beta+c},\end{aligned}$$

where the last step follows from fact B.1. ∎

The following two facts about the Poisson distribution will be useful.

Fact B.4 *Both the mean and the variance of a Poisson distributed variable S with Poisson parameter* θ *are* $\mathbb{E}[S\,|\theta] = \mathbb{V}(S\,|\theta) = \theta$.

Fact B.5 *The sum of N independently Poisson distributed variables with a common mean and variance* θ, $S_1 + ... + S_N$, *has a Poisson distribution with mean and variance* $N\theta$.

Proof. See DeGroot (1989, Ch. 5.4) for proofs of facts B.4 and B.5. ∎

B.3 *Ex ante* expected indirect utility

For CARA utility, the intertemporal expected utility function (4.1) is

$$U^i = -e^{-A(W_0^i - cN^i)}e^{A(b^i+Px^i)} - \delta e^{-ARb^i}\mathbb{E}^i\left[e^{-Ax^i\theta}\right] \tag{B.1}$$

by (4.2) and (4.3). Using $H^{i,*} \equiv \exp\left(-A\left[(1+R)b^{i,*} + Px^{i,*} - W_0^i + cN^i\right]\right)$, bond demand can be written as

$$b^{i,*} = \frac{1}{1+R}\left(W_0^i - cN^i - \frac{1}{A}\log H^{i,*} - Px^{i,*}\right)$$

and the portfolio value as

$$b^{i,*} + Px^{i,*} = \frac{1}{1+R}\left(W_0^i - cN^i - \frac{1}{A}\log H^{i,*} + RPx^{i,*}\right).$$

$H^{i,*}$ is certain and implicitly given by the first order condition (4.4). Using these facts in (B.1) yields

$$\begin{aligned}U^{i,*} &= -e^{-A\frac{R}{1+R}(W_0^i - cN^i)}e^{-\frac{1}{1+R}\log H^{i,*}}e^{A\frac{R}{1+R}Px^{i,*}}(1+\delta e^{\log H^{i,*}}e^{-Ax^{i,*}\theta}) \\ &= -\exp\left\{-A\frac{R}{1+R}(W_0^i - cN^i)\right\}\left(\frac{e^{ARPx^{i,*}}}{H^{i,*}}\right)^{\frac{1}{1+R}}\left(1+\frac{1}{R}\right),\end{aligned}$$

irrespective of the return (and signal) distribution. The second step follows by using the first order condition (4.4) to substitute for $H^{i,*}$. This yields indirect utility (4.11) in the text.

For a gamma distributed asset return and by (4.6), $ARPx^i = \beta^i \mathbb{E}^i[\theta - RP]$. In addition, $(H^i)^{-1} = \delta R \mathbb{E}^i[e^{-Ax^i\theta}] = \delta R(\frac{\beta^i}{\alpha^i} RP)^{\alpha^i}$ by (4.4) and fact B.2 (p. 261). This implies (4.12) in the text. To derive *ex ante* (expected indirect) utility, the following property of the gamma-Poisson conjugate pair is useful.

Fact B.6 *For two arbitrary constants B and G, N Poisson distributed signals $S_1, ..., S_N$ and a conjugate prior gamma distribution of their common mean θ, the following is true.*

$$\mathbb{E}_{prior}\left[(1+G)^{-B\cdot\sum_{n=1}^{N} s_n} \cdot e^{\frac{G}{1+G}B\cdot\sum_{n=1}^{N} s_n}\right]$$
$$= (1+G)^{\bar{\alpha}B} e^{-\bar{\alpha}\frac{G}{1+G}B}\left(\frac{1}{1+\left[(1+G)^B e^{-\frac{G}{1+G}B} - 1\right]\frac{\beta}{\bar{\beta}}}\right)^{\bar{\alpha}},$$

where $\bar{\alpha}$ and $\bar{\beta}$ are the parameters of the prior gamma distribution of θ, and $\beta = \bar{\beta} + N$ is the according parameter of the posterior distribution.

Proof. By the law of iterated expectations $\mathbb{E}_{prior}[\cdot] = \mathbb{E}_\theta[\mathbb{E}[\cdot|\theta]]$. The 'inner' expectation $\mathbb{E}[\cdot|\theta]$ is equal to

$$\begin{aligned}
\mathbb{E}[\cdot|\theta] &= \sum_{(\sum_{n=1}^{N} s_n)=0}^{\infty} (1+G)^{-B\sum_{n=1}^{N} s_n} e^{\frac{G}{1+G}B\sum_{n=1}^{N} s_n} f\left(\sum_{n=1}^{N} s_n\right) \\
&= \exp\left\{-N\theta\left(1-(1+G)^{-B}\exp\{\frac{G}{1+G}B\}\right)\right\},
\end{aligned}$$

because the sum $\sum_{n=1}^{N} s_n$ is Poisson distributed with mean $N\theta$ (fact B.5, p. 262). By fact 4.1 (p. 111), $N = \beta - \bar{\beta}$. Thus, by fact B.2 (p. 261),

$$\begin{aligned}
\mathbb{E}_{prior}[\cdot] &= \mathbb{E}_\theta\left[\exp\left\{-\theta\left(1-(1+G)^{-B}\exp\{\frac{G}{1+G}B\}\right)(\beta-\bar{\beta})\right\}\right] \\
&= \left(\frac{\bar{\beta}}{\bar{\beta}+\left(1-(1+G)^{-B}\exp\left(\frac{G}{1+G}B\right)\right)(\beta-\bar{\beta})}\right)^{\bar{\alpha}}.
\end{aligned}$$

Simplifying the denominator and factoring out $(1+G)^B e^{-\frac{G}{1+G}B}$ proves fact B.6. ■

With fact B.6 at hand, one can set $G \equiv (A\bar{x})/(I\beta)$ and $B \equiv \frac{1}{1+R}$ and the last step in (4.14) follows immediately.

B.4 Properties of the Potential Information Benefit Curve

Properties of the potential marginal benefit of information (the second term on the right hand side of (4.17)) are derived here. Ultimately, the following lemma will be established.

Lemma B.1 *The potential marginal benefit of information*

$$\frac{\bar{\alpha}}{\bar{\beta}} \frac{\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}} \left(1 - \frac{1}{1+R}\frac{\xi^2}{(1+\xi)^2}\right) - 1}{1 + \left(\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}} - 1\right)\frac{\bar{\xi}}{\xi}} \tag{B.2}$$

attains a strictly positive value for sufficiently large $\xi > 0$ and $R \in [0, \infty)$. In this positive range of ξ, the marginal benefit is strictly monotonically increasing in ξ and unbounded.

The proof of lemma B.1 proceeds in four steps. First, it is shown that the denominator

$$h(\xi) \equiv 1 + \left(\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}} - 1\right)\frac{\bar{\xi}}{\xi}. \tag{B.3}$$

is bounded below and above, and that it is strictly decreasing in ξ iff the numerator is strictly positive. So, the denominator cannot explode and tends to boost the marginal benefit higher once the potential benefit becomes positive. Second, the benefit reducing factor $\left(1 - [1/(1+R)]\,[\xi/(1+\xi)]^2\right)$ in the numerator

$$g(\xi) \equiv \left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}} \left(1 - \frac{1}{1+R}\frac{\xi^2}{(1+\xi)^2}\right) - 1 \tag{B.4}$$

is shown to be strictly decreasing in ξ but bounded below. Based on this second finding, it is shown third that the numerator $g(\xi)$ as a whole is strictly increasing in ξ for $\xi > \sqrt{1 + 1/R}$ and that it is not bounded above. So, the numerator boosts the marginal benefit higher and higher as ξ rises. Fourth, these facts together imply that the marginal benefit term (B.2) is strictly increasing in ξ in the positive range, and unbounded, and that it reaches this strictly positive range for $R \in [0, \infty)$.

Claim B.1 *The denominator $h(\xi)$ is bounded below by 1 and above by $1 + \bar{\xi}/e$ for $\xi > 0$ and $R \in [0, \infty)$, where $\bar{\xi}$ is given by (4.15) and e is Euler's number. $h(\xi)$ is strictly decreasing in ξ iff $g(\xi) > 0$.*

Proof. By L'Hôpital's rule, $\lim_{\xi \to 0} \left[(1+\xi)\exp(-\frac{\xi}{1+\xi})\right]^{\frac{1}{1+R}} / \xi - 1/\xi = 0$. This establishes the lower bound.

To find the behavior of $h(\xi)$ when ξ goes to infinity, observe that $h(\xi)$ grows most strongly in ξ if $R = 0$. So, consider the benchmark case of $R{=}0$ first. Then, $\lim_{\xi \to \infty} (\frac{1+\xi}{\xi})\exp(-\frac{\xi}{1+\xi}) - 1/\xi = 1/e$, where e is Euler's number. Note that the term $(\frac{1+\xi}{\xi})\exp(-\frac{\xi}{1+\xi}) - 1/\xi$ is strictly increasing in

ξ for $\xi > 0$ because, by the properties of the log function, $\log[(1+\xi)/(1+2\xi)] > -\xi/(1+\xi)$ for $\xi > 0$, so that the first derivative of the term is always positive. Hence, $h(\xi)$ cannot exceed a value of $1+\bar{\xi}/e$ at any $\xi > 0$ for $R = 0$. Since $h(\xi)$ is strictly decreasing in R, $\lim_{\xi\to\infty} h(\xi) \leq 1+\bar{\xi}/e$ must be true for any other $R \in [0,\infty)$. Hence, $1+\bar{\xi}/e$ is a global upper bound on $h(\xi)$ for any $R \in [0,\infty)$ and $\xi > 0$.

The first derivative of $h(\xi)$ is

$$\frac{\partial}{\partial\xi} h(\xi) = -\frac{\bar{\xi}}{\xi^2}\left(\left[(1+\xi)e^{-\frac{\xi}{1+\xi}}\right]^{\frac{1}{1+R}}\left(1-\frac{1}{1+R}\frac{\xi^2}{(1+\xi)^2}\right)-1\right).$$

It is strictly negative if and only if $g(\xi)$ is strictly positive. ■

Let's turn to the numerator.

Claim B.2 *For $R \in [0,\infty)$, the factor $\left(1-[1/(1+R)]\,[\xi/(1+\xi)]^2\right)$ in the numerator is strictly decreasing in ξ. For $R \in [0,\infty)$ and $\xi > 0$, $\left(1-[1/(1+R)]\,[\xi/(1+\xi)]^2\right) \in (0,1)$.*

Proof. Since $\xi/(1+\xi)$ is strictly increasing in ξ, the factor is strictly decreasing. That the factor cannot exceed unity is guaranteed by $R \geq 0$. Suppose the factor could fall below zero so that $\left(1-[1/(1+R)]\,[\xi/(1+\xi)]^2\right) < 0$ for some $\xi > 0$. Then $\xi/(1+\xi) > +\sqrt{1+R}$ must hold at that ξ. This implies, however, that $(1-\sqrt{1+R})\xi > \sqrt{1+R}$. Since $R \in [0,\infty)$ by assumption, this can never be the case. ■

With this result at hand, properties of the numerator as a whole can be established.

Claim B.3 *The numerator $g(\xi)$ is strictly decreasing in ξ iff $\xi \in (0, \sqrt{1+1/R})$ and strictly increasing iff $\xi \in (\sqrt{1+1/R}, \infty)$. In addition, $\lim_{\xi\to 0} g(\xi) = 0$ and $\lim_{\xi\to\infty} g(\xi) = +\infty$.*

Proof. Note that any function is strictly increasing if and only if a strictly monotone transformation of it is strictly increasing. So, the numerator $g(\xi)$ is strictly increasing in ξ iff $\log(1+g(\xi))$ is. Taking the first derivative of this function and simplifying establishes that

$$\frac{\partial}{\partial\xi}\log(1+g(\xi)) = \frac{\xi}{(1+\xi)^2} - \frac{\xi}{(1+\xi)^3}\frac{2}{1-\frac{1}{1+R}\frac{\xi^2}{(1+\xi)^2}} > 0 \text{ iff } \xi > \sqrt{1+\frac{1}{R}}.$$

This follows from the fact that the factor $\left(1-[1/(1+R)]\,[\xi/(1+\xi)]^2\right) > 0$ by claim B.2.

That $g(0) = 0$ is easily seen. To find the limit of $g(\xi)$ as ξ tends to infinity, suppose $g(\xi)$ was bounded above. Then, the strictly monotone transformation $\log(1+g(\xi))$ would have to be bounded above. However, $\lim_{\xi\to\infty}\log(1+\xi)-\xi/(1+\xi) = +\infty$ and $\lim_{\xi\to\infty}\left(1-[1/(1+R)]\,[\xi/(1+\xi)]^2\right) = 0$ by claim B.2. So, $g(\xi)$ cannot be bounded above. ■

Taken together, the three previous claims establish that the marginal benefit term (B.2) is strictly and unboundedly increasing in ξ as soon as ξ once reached the weakly positive range. The reason is a combination of the following three facts. First, the denominator is decreasing in the weakly positive range of (B.2) by claim B.1. Second, the benefit term (B.2) goes to zero as ξ does (by claims B.1 and B.3). Hence, third, if the benefit term reaches the weakly positive range, it must get there for some $\xi > \sqrt{1 + 1/R}$ because both the numerator (by claim B.3) and the denominator (by claim B.1) are strictly reducing (B.2) for lower ξ values.

Thus, for lemma B.1 to be proven, it only remains to be shown that the benefit term has at least one zero point for $\xi \in (0, \infty)$. We know that $\lim_{\xi \to \infty} g(\xi) = +\infty$ by claim B.3 and $\lim_{\xi \to \infty} h(\xi) \leq 1 + \bar{\xi}/e$ by claim B.1. Thus, $\lim_{\xi \to \infty} g(\xi)/h(\xi) = +\infty$ so that there must be at least one zero point of (B.2). By the same argument, the benefit term has at most one zero point for $\xi \in (0, \infty)$. This concludes the proof of lemma B.1.

Appendix C

Mathematical appendix to chapter 5

This appendix provides the proofs of propositions in chapter 5 and presents properties of jointly normally distributed variables that are invoked at various stages of chapter 5.

C.1 Properties of the normal distribution

A rational (Bayesian) investor updates her beliefs using the conditional normal distribution of the dividend given the signal and price realizations. Since signals and price are not conditionally independent, rational investors will make use of the following fact in general.

Fact C.1 *Consider a multivariate normal p.d.f.* $f\left((\theta;\mathbf{z}^T)\,|\mu,\Sigma\right)$ *with* $\mathbf{Z} = (Z_1,...,Z_K)^T$, $\mu \equiv (\bar{\mu}_\theta;\mathbb{E}[Z_1],...,\mathbb{E}[Z_K])^T$ *and*

$$\Sigma \equiv \begin{pmatrix} \bar{\tau}_\theta^2 & \mathbb{C}\mathsf{ov}(\theta.\mathbf{Z})^T \\ \mathbb{C}\mathsf{ov}(\theta.\mathbf{Z}) & \mathbb{C}\mathsf{ov}\left(\mathbf{Z}.\mathbf{Z}^T\right) \end{pmatrix}.$$

Then the conditional p.d.f. of θ*, given a vector* $\mathbf{z}$ *of realizations of* $\mathbf{Z}$ *is normal with*

$$f\Big(\theta \;\Big|\; \bar{\mu}_\theta + \mathbb{C}\mathsf{ov}(\theta.\mathbf{Z})^T\,\mathbb{C}\mathsf{ov}\left(\mathbf{Z}.\mathbf{Z}^T\right)^{-1}(\mathbf{z}-\mathbb{E}[z]),\; \left[\bar{\tau}_\theta^2 - \mathbb{C}\mathsf{ov}(\theta.\mathbf{Z})^T\,\mathbb{C}\mathsf{ov}\left(\mathbf{Z}.\mathbf{Z}^T\right)^{-1}\mathbb{C}\mathsf{ov}(\theta.\mathbf{Z})\right]^{-1}\Big).$$

Proof. See Raiffa and Schlaifer (1961, 8.2.1). ∎

Fact 5.1 (p. 145) is a special case of fact C.1 when all signals are conditionally independent.

Apart from this property, three further characteristics of the normal distribution are of use in the present framework.

Fact C.2 *For a normally distributed random variable* $z \sim \mathcal{N}\left(\mu, \sigma^2\right)$ *and an arbitrary constant* A*, the expected value of* $e^{-A \cdot z}$ *is*

$$\mathbb{E}\left[e^{-A \cdot z} \left| \mu, \sigma\right.\right] = \exp\left\{-A\mu + \frac{A^2}{2}\sigma^2\right\}$$

Fact C.3 *For a normally distributed random variable* $z \sim \mathcal{N}\left(\mu, \sigma^2\right)$ *and an arbitrary constant* A*, the expected value of* $z \cdot e^{-A \cdot z}$ *is*

$$\mathbb{E}\left[z e^{-A \cdot z} \left| \mu, \sigma\right.\right] = \left(\mu - A\sigma^2\right) \exp\left\{-A\mu + \frac{A^2}{2}\sigma^2\right\}.$$

Proof. Although fact C.2 is a well-known property, I will prove it again here since fact C.3 follows as a corollary. Note that

$$-\frac{1}{2}\left(\frac{z - \left(\mu - A\sigma^2\right)}{\sigma}\right)^2 = -A(z-\mu) - \frac{A^2\sigma^2}{2} - \frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2.$$

Thus,

$$\begin{aligned} \mathbb{E}\left[e^{-Az}\right] &= \int_{-\infty}^{\infty} e^{-Az} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} \mathrm{d}z \\ &= e^{-A\mu + \frac{A^2}{2}\sigma^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{z-(\mu - A\sigma^2)}{\sigma}\right)^2} \mathrm{d}z = e^{-A\mu + \frac{A^2}{2}\sigma^2}. \end{aligned}$$

This proves fact C.2. Similarly,

$$\begin{aligned} \mathbb{E}\left[z e^{-Az}\right] &= e^{-A\mu + \frac{A^2}{2}\sigma^2} \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{z-(\mu - A\sigma^2)}{\sigma}\right)^2} \mathrm{d}z \\ &= e^{-A\mu + \frac{A^2}{2}\sigma^2} \left[\mu - A\sigma^2\right], \end{aligned}$$

and fact C.3 follows. ■

Finally, the following fact is useful to derive *ex ante* utility in the case of partly informative prices.

Fact C.4 *For a normally distributed random variable* $z \sim \mathcal{N}\left(\mu, \sigma^2\right)$ *and three arbitrary constants* A, B, D*, the expected value of* $e^{-\frac{A}{2}(B + D\,z)^2}$ *is*

$$\mathbb{E}\left[e^{-\frac{A}{2}(B+D\,z)^2} \left| \mu, \sigma\right.\right] = \frac{1}{\sqrt{1 + A\,D^2\sigma^2}} \exp\left\{-\frac{A}{2}\frac{(B + D\mu)^2}{1 + A\,D^2\sigma^2}\right\}.$$

Proof. To derive this fact, consider the expectations of $e^{-A_1 z - A_2 z^2}$ for two arbitrary constants A_1, A_2. Note that

$$-\frac{1}{2}\left(\frac{z-\left[\mu-(1+2\frac{A_2}{A_1}\frac{\mu-A_1\sigma^2}{1+2A_2\sigma^2})A_1\sigma^2\right]}{\frac{\sigma}{\sqrt{1+2A_2\sigma^2}}}\right)^2$$

$$= -z(A_1+A_2 z)+\frac{\mu(A_1+A_2\mu)-\frac{c_1^2}{2}\sigma^2}{1+A_2\sigma^2}-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2.$$

Thus,

$$\mathbb{E}\left[e^{-A_1 z - A_2 z^2}\right] = \int_{-\infty}^{\infty} e^{-A_1 z - A_2 z^2}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2}\,\mathrm{d}z$$

$$= \frac{e^{-\frac{\mu(A_1+A_2\mu)-\frac{A_1^2}{2}\sigma^2}{1+A_2\sigma^2}}}{\sqrt{1+2A_2\sigma^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{1+2A_2\sigma^2}}}e^{-\frac{1}{2}\left(\frac{z-\left[\mu-(1+2\frac{A_2}{A_1}\frac{\mu-A_1\sigma^2}{1+2A_2\sigma^2})A_1\sigma^2\right]}{\frac{\sigma}{\sqrt{1+2A_2\sigma^2}}}\right)^2}\,\mathrm{d}z$$

$$= \frac{1}{\sqrt{1+2A_2\sigma^2}}\exp\left\{-\frac{\mu(A_1+A_2\mu)-\frac{(A_1)^2}{2}\sigma^2}{1+A_2\sigma^2}\right\}. \tag{C.1}$$

To arrive at fact C.4, observe that

$$\mathbb{E}\left[e^{-\frac{A}{2}(B+D\,z)^2}\right] = e^{-\frac{A}{2}B^2}\mathbb{E}\left[e^{-\frac{A}{2}(2BDz+D^2z^2)}\right].$$

Then defining $A_1 \equiv \frac{A}{2}2BD$ and $A_2 \equiv \frac{A}{2}D^2$, multiplying (C.1) by $e^{-\frac{A}{2}B^2}$, and collecting terms yields fact C.4. ∎

C.2 Posterior indirect expected utility

Using $H^i \equiv \exp\left(-A\left[(1+R)b^i + Px^i - W_0^i + F^i + cN^i\right]\right)$ and solving out for b^i yields demand for the bond

$$b^{i,*} = \frac{1}{1+R}\left(W_0^i - F^i - cN^i - Px^{i,*} - \frac{1}{A}\ln H^{i,*}\right). \tag{C.2}$$

For each unit of the risky asset, bond demand is adjusted by a factor of $P/(1+R)$ to achieve tomorrow's desired consumption level.

To derive indirect utility (5.10) in the text, note that (5.1) simplifies to

$$U^i = -e^{-A(W_0^i - F^i - cN^i)}e^{A(b^i+Px^i)} - \delta e^{-ARb^i}\mathbb{E}^i\left[e^{-Ax^i\theta}\right] \tag{C.3}$$

for CARA utility. By (C.2) (which holds for CARA utility irrespective of the risky asset's distribution), we can write

$$b^{i,*} + Px^{i,*} = \frac{1}{1+R}\left(W_0^i - F^i - cN^i - \frac{1}{A}\ln H^{i,*} + RPx^{i,*}\right),$$

where $H^{i,*}$ is certain and implicitly given by the first order condition (5.6). Using the above fact and (C.2) in (C.3) yields

$$\begin{aligned} U^{i,*} &= -e^{-A\frac{R}{1+R}(W_0^i - F^i - cN^i)} e^{-\frac{1}{1+R}\ln H^{i,*}} e^{A\frac{R}{1+R}Px^{i,*}}(1+\delta e^{\ln H^{i,*}} e^{-Ax^{i,*}\theta}) \\ &= -\exp\left\{-A\frac{R}{1+R}(W_0^i - F^i - cN^i)\right\}\left(\frac{e^{ARPx^{i,*}}}{H^{i,*}}\right)^{\frac{1}{1+R}}\left(1+\frac{1}{R}\right). \end{aligned}$$

The second step follows by using the first order condition (5.6) to substitute for $H^{i,*}$. This establishes (5.10) in the text.

C.3 News watchers' *ex ante* distribution

Take a news watcher's point of view. Given any choice of N, the *ex ante* joint normal distribution of θ, N signals, and RP, that is the *ex ante* distribution of $(\theta; S_1, ..., S_N; RP)^T$ has a vector of means

$$\bar{\mu}^{NW} = (\bar{\mu}_\theta; \bar{\mu}_\theta, ..., \bar{\mu}_\theta; \pi_0 + \pi_S N\bar{\mu}_\theta - \pi_X \bar{x})^T$$

and an $(N+2)\times(N+2)$ variance-covariance matrix

$$\bar{\Sigma}^{NW} = \begin{pmatrix} \bar{\tau}_\theta^2 & \bar{\tau}_\theta^2 \cdot \iota_N^T & \pi_S N \bar{\tau}_\theta^2 \\ \bar{\tau}_\theta^2 \cdot \iota_N & \mathbb{C}\mathsf{ov}(\mathbf{S}.\mathbf{S}^T)_N & \pi_S(N\bar{\tau}_\theta^2 + \sigma_S^2)\cdot \iota_N \\ \pi_S N\bar{\tau}_\theta^2 & \pi_S(N\bar{\tau}_\theta^2 + \sigma_S^2)\cdot \iota_N^T & \pi_S^2 N(N\bar{\tau}_\theta^2 + \sigma_S^2) + \pi_X^2\omega_X^2 \end{pmatrix}.$$

$\mathbf{S} = (S_1, ..., S_N)^T$ is the vector of N signals, ι_N denotes an N vector of ones, and

$$\mathbb{C}\mathsf{ov}(\mathbf{S}.\mathbf{S}^T)_N = \begin{pmatrix} \bar{\tau}_\theta^2 + \sigma_S^2 & \bar{\tau}_\theta^2 & \cdots & \bar{\tau}_\theta^2 \\ \bar{\tau}_\theta^2 & \bar{\tau}_\theta^2 + \sigma_S^2 & & \bar{\tau}_\theta^2 \\ \vdots & & \ddots & \vdots \\ \bar{\tau}_\theta^2 & \cdots & & \bar{\tau}_\theta^2 + \sigma_S^2 \end{pmatrix}.$$

After observing signal realizations $(s_1, ..., s_N)$ and RP, news watchers apply fact C.1 to this *ex ante* joint normal distribution and obtain a posterior normal distribution of the dividend with conditional mean

$$\mathbb{E}\left[\theta\,|RP; s_1, ..., s_N; \lambda, N\right] = \mu^{NW} = m_0^{NW} + m_S^{NW}\sum_{j=1}^{N} s_j + m_{RP}^{NW} RP \qquad \text{(C.4)}$$

and conditional variance $\mathbb{V}(\theta \mid RP; s_1, ..., s_N; \lambda, N) = (\tau^{NW})^2$, where

$$\begin{aligned} m_0^{NW} &= \frac{\sigma_S^2 \bar{\mu}_\theta}{\sigma_S^2 + \bar{\tau}_\theta^2 N}, && \text{(C.5)} \\ m_S^{NW} &= \frac{\bar{\tau}_\theta^2}{\sigma_S^2 + \bar{\tau}_\theta^2 N}, && \text{(C.6)} \\ m_{RP}^{NW} &= 0, && \text{(C.7)} \\ (\tau^{NW})^2 &= \frac{\sigma_S^2 \bar{\tau}_\theta^2}{\sigma_S^2 + \bar{\tau}_\theta^2 N}. && \text{(C.8)} \end{aligned}$$

This is precisely what has been stated in fact 5.1 (p. 145) before.

C.4 Two-group financial market equilibrium

A two-group financial market equilibrium is given by matching the coefficients π_0, π_S, π_X in equation (5.16) with the according terms in (5.25). Defining

$$u \equiv \frac{1}{\bar{\tau}_\theta^2} + \left[(1-\lambda)\frac{\pi_S(\pi_S N - 1)}{\pi_S^2 N \sigma_S^2 + \pi_X^2 \omega_X^2} + \lambda \frac{1}{\sigma_S^2}\right] N \tag{C.9}$$

and matching coefficients π_0, π_S, π_X yields

$$\begin{aligned} \pi_0 &= \frac{1}{u}\left(\frac{\bar{\mu}_\theta}{\bar{\tau}_\theta^2} - (1-\lambda)\frac{\pi_S N(\pi_0 - \pi_X \bar{x})}{\pi_S^2 N \sigma_S^2 + \pi_X^2 \omega_X^2}\right), && \text{(C.10)} \\ \pi_S &= \frac{1}{u}\frac{\lambda}{\sigma_S^2}, && \text{(C.11)} \\ \pi_X &= \frac{1}{u}\frac{A}{I}. && \text{(C.12)} \end{aligned}$$

Plugging (C.11) and (C.12) into (C.9) and simplifying shows that (C.9) is a linear equation indeed. In general, if there are I^{NW} (groups) of investors who acquire a strictly positive number of signals, u is a polynomial of order $1 + 2I^{NW}$ (see appendix C.9). Here, however, u has the unique solution

$$u = \frac{1}{\bar{\tau}_\theta^2} + \left(\frac{1}{\sigma_S^2} - \frac{1}{\bar{\tau}_\theta^2}\frac{(1-\lambda)I^2}{\lambda I \cdot NI + A^2 \sigma_S^2 \omega_X^2}\right)\lambda N.$$

Hence,

$$\begin{aligned} \pi_0 &= \frac{\left[(\lambda I)^2 N + A^2 \sigma_S^2 \omega_X^2\right] \sigma_S^2 \cdot \bar{\mu}_\theta + (\lambda I)(1-\lambda) N \sigma_S^2 \bar{\tau}_\theta^2 \cdot A\bar{x}}{(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2 \sigma_S^2 \omega_X^2 (\sigma_S^2 + \lambda N \bar{\tau}_\theta^2)}, && \text{(C.13)} \\ \pi_S &= \frac{1}{\frac{1}{\bar{\tau}_\theta^2} + \left(\frac{1}{\sigma_S^2} - \frac{1}{\bar{\tau}_\theta^2}\frac{(1-\lambda)I^2}{\lambda I \cdot NI + A^2 \sigma_S^2 \omega_X^2}\right)\lambda N}\frac{\lambda}{\sigma_S^2}, && \text{(C.14)} \\ \pi_X &= \frac{1}{\frac{1}{\bar{\tau}_\theta^2} + \left(\frac{1}{\sigma_S^2} - \frac{1}{\bar{\tau}_\theta^2}\frac{(1-\lambda)I^2}{\lambda I \cdot NI + A^2 \sigma_S^2 \omega_X^2}\right)\lambda N}\frac{A}{I}. && \text{(C.15)} \end{aligned}$$

The key term for both investors is $(\mu^i - RP)/\tau^i = \tau^i (\mu^i - RP)/(\tau^i)^2$. To solve for the according price watcher term, first plug (C.13) through (C.15) into m_0^{PW} (5.18), m_{RP}^{PW} (5.19), and

$(\tau^{PW})^2$ (5.20). This yields $(\tau^{PW})^2$. Then plug the solutions for m_0^{PW}, m_{RP}^{PW}, and $(\tau^{PW})^2$ along with the solution for the opportunity cost RP (5.16) into $(\mu^{PW} - RP)/(\tau^{PW})^2$. Collecting terms and simplifying yields

$$\frac{\mu^{PW} - RP}{(\tau^{PW})^2} = \frac{A}{I} \frac{1}{(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)} \cdot \left((\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) \cdot \bar{x} - \lambda I A \sigma_S^2 \omega_X^2 \cdot \sum_{j=1}^{N} (S_j - \bar{\mu}_\theta) + A^2 \sigma_S^4 \omega_X^2 \cdot X \right) \tag{C.16}$$

and

$$(\tau^{PW})^2 = \frac{\left[(\lambda I)^2 N\sigma_S^2 + A^2\sigma_S^4\omega_X^2\right] \bar{\tau}_\theta^2}{(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2} \tag{C.17}$$

for price watchers. Similarly, using (C.13) through (C.15) for news watchers yields

$$\frac{\mu^{NW} - RP}{(\tau^{NW})^2} = \frac{A}{I} \frac{1}{(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)} \cdot \Bigg(- \lambda I(1-\lambda) I N(\sigma_S^2 + N\bar{\tau}_\theta^2) \cdot \bar{x} + (1-\lambda) I A \sigma_S^2 \omega_X^2 \cdot \sum_{j=1}^{N} (S_j - \bar{\mu}_\theta) + \left(I^2 \lambda N + A^2\sigma_S^4\omega_X^2\right)(\sigma_S^2 + N\bar{\tau}_\theta^2) \cdot X \Bigg) \tag{C.18}$$

by (5.22), (5.23), and (5.24) along with (5.16), while

$$(\tau^{NW})^2 = \frac{\sigma_S^2 \bar{\tau}_\theta^2}{\sigma_S^2 + \bar{\tau}_\theta^2 N} \tag{C.19}$$

as given in (5.24).

C.5 Moments of key term

In subsequent analysis, the *ex ante* moments of the key term $\tau^i\,(\mu^i - RP)/(\tau^i)^2$ will be of most interest. Since $(\tau^i)^2$ is certain, and both μ^i and RP are normally distributed from a *ex ante* perspective, $(\mu^i - RP)/(\tau^i)^2$ is normally distributed. To derive the moments, start with price watchers. Taking expectations and the variance of (C.16), one finds

$$\mathbb{E}_{pre}^{PW}\left[\frac{\mu^{PW} - RP}{(\tau^{PW})^2}\right] = \frac{A}{I} \frac{\left[(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2\right]\, \bar{x}}{(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)}, \tag{C.20}$$

$$\mathbb{V}_{pre}^{PW}\left(\frac{\mu^{PW} - RP}{(\tau^{PW})^2}\right) = \frac{A^2}{I^2} \frac{\left[(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2\right]\, A^2\sigma_S^4\omega_X^4}{\left[(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right]^2} \tag{C.21}$$

for price watchers. Their $(\tau^{PW})^2$ is

$$(\tau^{PW})^2 = \frac{\left[(\lambda I)^2 N\sigma_S^2 + A^2\sigma_S^4\omega_X^2\right] \bar{\tau}_\theta^2}{(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2} \tag{C.22}$$

by (C.17).

Similarly, taking expectations and the variance of (C.18) for news watchers, one finds

$$\mathbb{E}^{NW}_{pre}\left[\frac{\mu^{NW}-RP}{(\tau^{NW})^2}\right]=\frac{A}{I}\frac{\left[(\lambda I)^2N+A^2\sigma_S^4\omega_X^2\right](\sigma_S^2+N\bar{\tau}_\theta^2)\,\bar{x}}{(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)}, \tag{C.23}$$

$$\begin{aligned}\mathbb{V}^{NW}_{pre}\left(\frac{\mu^{NW}-RP}{(\tau^{NW})^2}\right)=\frac{A^2}{I^2}&\frac{\omega_X^2(\sigma_S^2+N\bar{\tau}_\theta^2)}{\left[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\right]^2}\\ \cdot\Bigg(&\lambda^2I^4N^2(\sigma_S^2+N\bar{\tau}_\theta^2)+I^2NA^2\sigma_S^2\omega_X^2\left((1+\lambda^2)\sigma_S^2+2\lambda N\bar{\tau}_\theta^2\right)\\ &+A^4\sigma_S^4\omega_X^4(\sigma_S^2+N\bar{\tau}_\theta^2)\Bigg).\end{aligned} \tag{C.24}$$

Their $(\tau^{NW})^2$ is

$$(\tau^{NW})^2=\frac{\sigma_S^2\bar{\tau}_\theta^2}{\sigma_S^2+N\bar{\tau}_\theta^2} \tag{C.25}$$

by (C.19).

If $N=0$, news watchers' terms (C.23), (C.24), and (C.19) coincide with the respective terms for price watchers (C.20), (C.21), and (C.17), as it should be.

C.6 Two-group information market equilibrium

Written out, the derivative of *ex ante* utility (5.27) with respect to N, given λ, is

$$\begin{aligned}\frac{1+R}{\mathbb{E}^i_{pre}\left[U^{i,*}\right]}\frac{\partial\mathbb{E}^i_{pre}\left[U^{i,*}\right]}{\partial N^i}=&\ -A\,R\,c\\ &+\frac{1}{N}\frac{\left(\mathbb{E}^i_{pre}\left[\frac{\mu^i-RP}{(\tau^i)^2}\right]\right)^2}{\frac{1}{(\tau^i)^2}+\frac{1}{1+R}\mathbb{V}^i_{pre}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}\left(\varepsilon^i_{\tau^2,N}+\varepsilon^i_{\mathbb{E},N}\right)\\ &-\frac{1}{N}\frac{\mathbb{V}^i_{pre}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}{\frac{1}{(\tau^i)^2}+\frac{1}{1+R}\mathbb{V}^i_{pre}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}\left(\varepsilon^i_{\tau^2,N}+\tfrac{1}{2}\varepsilon^i_{\mathbb{V},N}\right)\\ &\qquad\cdot\frac{1}{1+R}\frac{\left(\mathbb{E}^i_{pre}\left[\frac{\mu^i-RP}{(\tau^i)^2}\right]\right)^2-\frac{1+R}{(\tau^i)^2}-\mathbb{V}^i_{pre}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)}{\frac{1}{(\tau^i)^2}+\frac{1}{1+R}\mathbb{V}^i_{pre}\left(\frac{\mu^i-RP}{(\tau^i)^2}\right)},\end{aligned} \tag{C.26}$$

where $\varepsilon^i_{y,N}$ denotes the elasticity of y with respect to N. In the text, this condition has been given a nicer look be defining E^i, V^i, and Δ^i accordingly.

The terms in condition (C.26) can all be evaluated in closed-form using the moments of the key term $\tau^i\,(\mu^i-RP)/(\tau^i)^2$ as derived in the previous appendix C.5. Again, start with price watchers. Differentiating the moments (C.20) and (C.21) and $(\tau^{PW})^2$ with respect to N yields the elasticities

$$\begin{aligned}
\varepsilon^{PW}_{\mathbb{E},N} &= \frac{A^2\sigma_S^2\bar{\tau}_\theta^2\omega_X^2\,\lambda N}{[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^4\omega_X^2]} \\
&\quad\cdot\frac{(\lambda I)^2N^2\bar{\tau}_\theta^2-A^2\sigma_S^4\omega_X^2}{[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)]}, \qquad \text{(C.27)}\\
\varepsilon^{PW}_{\mathbb{V},N} &= -\frac{\lambda N}{[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^4\omega_X^2]} \\
&\quad \frac{1}{[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)]} \\
&\quad\cdot\Bigg(\lambda^3I^4N(\sigma_S^2+N\bar{\tau}_\theta^2)(\sigma_S^2+2N\bar{\tau}_\theta^2) \\
&\quad +A^2\sigma_S^4\omega_X^2\lambda I^2\left(\sigma_S^2+(2+\lambda)N\bar{\tau}_\theta^2\right)+2A^4\sigma_S^6\bar{\tau}_\theta^2\omega_X^4\Bigg), \qquad \text{(C.28)}\\
\varepsilon^{PW}_{\tau^2,N} &= -\frac{(\lambda I)^2N^2\bar{\tau}_\theta^2\;[(\lambda I)^2N+2A^2\sigma_S^2\omega_X^2]}{[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^4\omega_X^2]\;[(\lambda I)^2N+A^2\sigma_S^2\omega_X^2]}. \qquad \text{(C.29)}
\end{aligned}$$

Differentiating news watchers' moments (C.23) and (C.24) and $(\tau^{NW})^2$ with respect to N yields the elasticities

$$\begin{aligned}
\varepsilon^{NW}_{\mathbb{E},N} &= -\frac{A^2\sigma_S^2\bar{\tau}_\theta^2\omega_X^2\,(1-\lambda)N}{[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^4\omega_X^2]} \\
&\quad\cdot\frac{(\lambda I)^2N^2\bar{\tau}_\theta^2-A^2\sigma_S^4\omega_X^2}{[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)]}, \qquad \text{(C.30)}\\
\varepsilon^{NW}_{\mathbb{V},N} &= -\frac{1}{v}\,\frac{A^2\sigma_S^4\omega_X^2\,(1-\lambda)N}{(\sigma_S^2+N\bar{\tau}_\theta^2)\,[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)]} \\
&\quad\cdot\Bigg((\lambda I)^2(1+\lambda)I^2N(\sigma_S^2+N\bar{\tau}_\theta^2)(\sigma_S^2+2N\bar{\tau}_\theta^2) \\
&\quad +A^2\sigma_S^2\omega_X^2I^2\left[(1+\lambda)\sigma_S^4+(2+\lambda(5+\lambda))\,N\sigma_S^2\bar{\tau}_\theta^2+6\lambda N^2\bar{\tau}_\theta^4)\right] \\
&\quad +2A^4\sigma_S^4\bar{\tau}_\theta^2\omega_X^4(\sigma_S^2+N\bar{\tau}_\theta^2)\Bigg), \qquad \text{(C.31)}\\
\varepsilon^{NW}_{\tau^2,N} &= -\frac{\bar{\tau}_\theta^2N}{\sigma_S^2+\bar{\tau}_\theta^2N}, \qquad \text{(C.32)}
\end{aligned}$$

where

$$v\equiv[\lambda^2I^3N^2(\sigma_S^2+N\bar{\tau}_\theta^2)+I^2NA^2\sigma_S^2\omega_X^2\left((1+\lambda^2)\sigma_S^2+2\lambda N\bar{\tau}_\theta^2)\right) +A^4\sigma_S^4\omega_X^4(\sigma_S^2+N\bar{\tau}_\theta^2)].$$

With all these results at hand, we can evaluate (C.26). Take the E^i-terms first. For a price watcher

$$\frac{1}{N}\,E^{PW}\cdot\left[\varepsilon^{PW}_{\tau^2,N}+\varepsilon^{PW}_{\mathbb{E},N}\right]=$$
$$-(1+R)\,\lambda\,\bar{x}^2A^2\sigma_S^2\bar{\tau}_\theta^4\left(I^2N\lambda^2(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^4\omega_X^2\right)$$
$$\cdot(\lambda^3I^4N^2+2\lambda I^2NA^2\sigma_S^2\omega_X^2+A^4\sigma_S^4\omega_X^4)\Bigg/$$
$$\Bigg(\left[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\right]$$
$$\cdot\Big[(1+R)\lambda^4I^6N^2(\sigma_S^2+N\bar{\tau}_\theta^2)^2+2(1+R)\lambda^2I^4NA^2\sigma_S^2\omega_X^2$$
$$\cdot(\sigma_S^2+N\bar{\tau}_\theta^2)(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)+I^2A^4\sigma_S^4\omega_X^4\Big((1+R)\sigma_S^4+A^6\sigma_S^8\bar{\tau}_\theta^2\omega_X^6$$
$$+\lambda N\sigma_S^2\bar{\tau}_\theta^2(2(1+R)+\lambda)+(1+R)\lambda^2N^2\bar{\tau}_\theta^4\Big)\Big]\Bigg)<0, \tag{C.33}$$

whereas for a news watcher

$$\frac{1}{N}\,E^{NW}\cdot\left[\varepsilon^{NW}_{\tau^2,N}+\varepsilon^{NW}_{\mathbb{E},N}\right]=$$
$$-(1+R)\,\lambda\,\bar{x}^2A^2\sigma_S^2\bar{\tau}_\theta^4(\sigma_S^2+N\bar{\tau}_\theta^2)(I^2N\lambda^2+A^2\sigma_S^2\omega_X^2)$$
$$\cdot(\lambda^3I^4N^2+2\lambda I^2NA^2\sigma_S^2\omega_X^2+A^4\sigma_S^4\omega_X^4)\Bigg/$$
$$\Bigg(\left[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\right]$$
$$\cdot\Big[(1+R)\lambda^4I^6N^2(\sigma_S^2+N\bar{\tau}_\theta^2)^2$$
$$+\lambda^2I^4NA^2\sigma_S^2\omega_X^2(\sigma_S^2+N\bar{\tau}_\theta^2)\left[2(1+R)\sigma_S^2+N\left(1+2(1+R)\lambda\right)\right]$$
$$+I^2A^4\sigma_S^4\omega_X^2\Big((1+R)\sigma_S^4+N\sigma_S^2\bar{\tau}_\theta^2\left(1+\lambda(2(1+R)+\lambda)\right)$$
$$+\lambda N^2\bar{\tau}_\theta^4(2+\lambda(1+R))\Big)+A^6\sigma_S^6\bar{\tau}_\theta^2\omega_X^2(\sigma_S^2+N\bar{\tau}_\theta^2)\Big]\Bigg)<0. \tag{C.34}$$

These and the following terms have been calculated and simplified using Mathematica 4.

Now consider the V^i-terms. For a price watcher

$$
\begin{aligned}
&\frac{1}{N}\, V^{PW} \cdot \left[\varepsilon^{PW}_{\tau^2,N} + \tfrac{1}{2}\varepsilon^{PW}_{\mathbb{V},N}\right] = \\
&\Bigg(\lambda\left((\lambda I)^2 N A^4 \sigma_S^4 \omega_X^2(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^6\sigma_S^8\omega_X^6\right) \\
&\quad \Big[\lambda^5 I^6 N^2(\sigma_S^2 + N\bar{\tau}_\theta^2)(\sigma_S^2 + 4N\bar{\tau}_\theta^2) + \lambda^3 I^4 N A^2\sigma_S^2\omega_X^2 \\
&\qquad \cdot\Big(2\sigma_S^2 + N(11+\lambda)\sigma_S^2\bar{\tau}_\theta^2 + 2N^2(3+\lambda)\bar{\tau}_\theta^4\Big) + \lambda I^2 A^4\sigma_S^4\omega_X^4 \\
&\qquad \cdot\Big(\sigma_S^4 + 3N\,(2+\lambda)\,\sigma_S^2\bar{\tau}_\theta^2 + 4\lambda N^2\bar{\tau}_\theta^4\Big) + 2A^6\sigma_S^8\bar{\tau}_\theta^2\omega_X^6\Big]\Bigg) \Bigg/ \\
&\Bigg(2\left((\lambda I)^2 N + A^2\sigma_S^2\omega_X^2\right)\left((\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2\right) \\
&\quad \left((\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right) \\
&\quad \left[\lambda^2 I^3 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + I A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right]^2 \\
&\quad \Bigg[\frac{1}{\bar{\tau}_\theta^2} + \frac{(\lambda I)^2 N^2}{(\lambda I)^2 N^2\sigma_S^2 + A^2\sigma_S^4\omega_X^2} \\
&\qquad + \left((\lambda I)^2 N A^4\sigma_S^4\omega_X^4(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^6\sigma_S^8\omega_X^6\right) \Big/ \Big[(1+R) \\
&\qquad\quad \cdot\Big(\lambda^2 I^3 N(\sigma_S^2 + N\bar{\tau}_\theta^2) I A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\Big)^2\Big]\Bigg]\Bigg) > 0,
\end{aligned}
\tag{C.35}
$$

and for a news watcher

$$
\begin{aligned}
&\frac{1}{N}\, V^{NW} \cdot \left[\varepsilon^{NW}_{\tau^2,N} + \tfrac{1}{2}\varepsilon^{NW}_{\mathbb{V},N}\right] = \\
&\Bigg((1+R)A^2\sigma_S^2\bar{\tau}_\theta^2)\omega_X^2 \\
&\quad \Big[2\lambda^4 I^6 N^3\bar{\tau}_\theta^2(\sigma_S^2 + N\bar{\tau}_\theta^2)^2 - \lambda^2 I^4 N A^2\sigma_S^2(\sigma_S^2 + N\bar{\tau}_\theta^2) \\
&\qquad \cdot\Big((1-\lambda^2)\sigma_S^4 - 2N(1+2\lambda^2)\sigma_S^2\bar{\tau}_\theta^2 - 6\lambda N^2\bar{\tau}_\theta^4\Big) - I^2 A^4\sigma_S^4\omega_X^4 \\
&\qquad \cdot\Big((1-\lambda^2)\sigma_S^6 + \lambda N\,(3-\lambda(8+\lambda))\,\sigma_S^4\bar{\tau}_\theta^2 - 2\lambda^2 N^2(5+\lambda)\sigma^2\bar{\tau}_\theta^4 \\
&\qquad\quad -6\lambda^2 N^3\bar{\tau}_\theta^6\Big) + 2\lambda A^6\sigma_S^6\bar{\tau}_\theta^2\omega_X^6(\sigma_S^2 + N\bar{\tau}_\theta^2)^2\Big]\Bigg) \Bigg/
\end{aligned}
$$

$$\Bigg(2(\sigma_S^2 + N\bar{\tau}_\theta^2)\left((\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right)$$

$$\Big[(1+R)\lambda^4 I^6 N^2(\sigma_S^2 + N\bar{\tau}_\theta^2)^2$$

$$+\lambda^2 I^4 N A^2\sigma_S^2\omega_X^2(\sigma_S^2 + N\bar{\tau}_\theta^2)\left(2(1+R)\sigma_S^2 N(1+2(1+R)\lambda)\right)$$

$$+I^2 A^4\sigma_S^2\omega_X^4\Big((1+R)\sigma^4 + N\sigma_S^2\bar{\tau}_\theta^2\left(1+\lambda(2(1+R)+\lambda)\right)$$

$$+\lambda N^2\bar{\tau}_\theta^2(2+(1+R)\lambda)\Big) + A^6\sigma_S^6\bar{\tau}_\theta^2\omega_X^6(\sigma_S^2 + N\bar{\tau}_\theta^2)\Big]\Bigg). \tag{C.36}$$

Finally, take the Δ^i terms. For a price watcher

$$\frac{1}{1+R}\Delta^{PW} =$$

$$-\Bigg(R + \Big[\left((\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2\right)$$

$$\Big(\lambda^2 I^4 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^4\sigma_S^4\bar{\tau}_\theta^2\omega_X^2(\omega_X^2 - \bar{x}^2)$$

$$+I^2 A^2\sigma_S^2\left(\sigma_S^2\omega_X^2 + \lambda N\bar{\tau}_\theta^2(2\omega_X^2 - \lambda\bar{x}^2)\right)\Big)\Big]\Big/$$

$$\left[\lambda^2 I^3 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + IA\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right]^2\Bigg)\Bigg/$$

$$\left(1 + R + \frac{A^4\sigma_S^6\bar{\tau}_\theta^2\omega_X^2\left((\lambda I)^2 N + A^2\sigma^2\omega_X^2\right)}{\left[\lambda^2 I^3 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + IA^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right]^2}\right), \tag{C.37}$$

while for a news watcher

$$\frac{1}{1+R}\Delta^{NW} =$$

$$-\Bigg(R + \Big[(\sigma_S^2 + N\bar{\tau}_\theta^2)\left((\lambda I)^2 N + A^2\sigma_S^2\omega_X^2\right)\Big(\lambda^2 I^4 N(\sigma_S^2 + N\bar{\tau}_\theta^2)$$

$$+A^4\sigma_S^4\bar{\tau}_\theta^2\omega_X^2(\omega_X^2 - \bar{x}^2) + I^2 A^2\sigma_S^2\left(\sigma_S^2\omega_X^2 + \lambda N\bar{\tau}_\theta^2(2\omega_X^2 - \lambda\bar{x}^2)\right)\Big)\Big]\Big/$$

$$\left[\lambda^2 I^3 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + IA\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right]^2\Bigg)\Bigg/$$

$$\Bigg(R + \Big[(\sigma_S^2 + N\bar{\tau}_\theta^2)\left((\lambda I)^2 N + A^2\sigma_S^2\omega_X^2\right)\Big(\lambda^2 I^4 N(\sigma_S^2 + N\bar{\tau}_\theta^2)$$

$$+I^2 A^2\sigma_S^2\omega_X^2\left(\sigma_S^2 + 2\lambda N\bar{\tau}_\theta^2\right) + A^4\sigma_S^4\bar{\tau}_\theta^2\omega_X^4\Big)\Big]\Big/$$

$$\left[\lambda^2 I^3 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + IA\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right]^2\Bigg). \tag{C.38}$$

The relationships in the last row of table 5.2 (p. 166) follow from

$$
\begin{aligned}
&(\tau^{PW})^2\left(\mathbb{E}^{PW}_{pre}\left[\frac{\mu^{PW}-RP}{(\tau^{PW})^2}\right]\right)^2-(\tau^{PW})^2\mathbb{V}^{PW}_{pre}\left[\frac{\mu^{PW}-RP}{(\tau^{PW})^2}\right]<1+R\\
\Leftrightarrow\quad &\frac{\bar{x}^2}{I^2}<\frac{1+R}{A^2\sigma_S^2\bar{\tau}_\theta^2}\frac{\left[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\right]^2}{((\lambda I)^2N+A^2\sigma_S^2\omega_X^2)\,((\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^4\omega_X^2)}\\
&\qquad+\frac{\omega_X^2}{I^2}\frac{A^2\sigma_S^4\omega_X^2}{(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^4\omega_X^2}\;\equiv\frac{(\bar{x}_c^{\Delta,PW})^2}{I^2}
\end{aligned}
\tag{C.39}
$$

by (C.20) and (C.21), and from

$$
\begin{aligned}
&(\tau^{NW})^2\left(\mathbb{E}^{NW}_{pre}\left[\frac{\mu^{NW}-RP}{(\tau^{NW})^2}\right]\right)^2-(\tau^{NW})^2\mathbb{V}^{NW}_{pre}\left[\frac{\mu^{NW}-RP}{(\tau^{NW})^2}\right]<1+R\\
\Leftrightarrow\quad &\frac{\bar{x}^2}{I^2}<\frac{1+R}{A^2\sigma_S^2\bar{\tau}_\theta^2}\frac{\left[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\right]^2}{((\lambda I)^2N+A^2\sigma_S^2\omega_X^2)^2\,(\sigma_S^2+N\bar{\tau}_\theta^2)}\\
&\qquad+\frac{\omega_X^2}{I^2}\left(1+\frac{\left((\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2\left[(1+\lambda)\sigma_S^2+2\lambda N\bar{\tau}_\theta^2\right]\right)}{(\sigma_S^2+N\bar{\tau}_\theta^2)\,((\lambda I)^2N+A^2\sigma_S^2\omega_X^2)^2}\right.\\
&\qquad\qquad\left.\cdot(1-\lambda)I^2N\right)\;\equiv\frac{(\bar{x}_c^{\Delta,NW})^2}{I^2}
\end{aligned}
\tag{C.40}
$$

by (C.23) and (C.24).

Similarly, the relationships in the middle column of table 5.2 (p.166, fourth and sixth row) can be inferred from

$$
\begin{aligned}
&E^{NW}\cdot\left[\varepsilon^{NW}_{\tau^2,N}+\varepsilon^{NW}_{\mathbb{E},N}\right]-E^{PW}\cdot\left[\varepsilon^{PW}_{\tau^2,N}+\varepsilon^{PW}_{\mathbb{E},N}\right]=\\
&-\Bigg(R(1+R)\lambda I^2N\bar{x}^2A^4\sigma_S^4\bar{\tau}_\theta^6\omega_X^2\\
&\quad\cdot\left((\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\right)\\
&\quad\cdot\left(\lambda^3I^4N^2+2\lambda I^2NA^2\sigma_S^2\omega_X^2A^4\sigma_S^4\omega_X^4\right)\Bigg)\Bigg/\\
&\quad\Bigg(\Big[(1+R)\lambda^4I^6N^2(\sigma_S^2+N\bar{\tau}_\theta^2)^2\\
&\qquad+2(1+R)\lambda^2I^4NA^2\sigma_S^2\omega_X^2(\sigma_S^2+N\bar{\tau}_\theta^2)(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\\
&\qquad+I^2A^4\sigma_S^4\omega_X^4\Big((1+R)\sigma_S^4+\lambda N\left(2(1+R)+\lambda\right)\sigma_S^2\bar{\tau}_\theta^2\\
&\qquad\quad+(1+R)\lambda^2N^2\bar{\tau}_\theta^4\Big)+A^6\sigma_S^8\bar{\tau}_\theta^2\omega_X^6\Big]\\
&\quad\cdot\Big[(1+R)\lambda^4I^6N^2(\sigma_S^2+N\bar{\tau}_\theta^2)^2+\lambda^2I^4NA^2\sigma_S^2\omega_X^2\\
&\qquad(\sigma_S^2+N\bar{\tau}_\theta^2)\left(2(1+R)\sigma_S^2+N\left(1+2(1+R)\lambda\right)\bar{\tau}_\theta^2\right)\\
&\qquad+I^2A^4\sigma_S^4\omega_X^4\Big((1+R)\sigma_S^4+N\left(1+\lambda\left(2(1+R)+\lambda\right)\right)\sigma_S^2\bar{\tau}_\theta^2\\
&\qquad\quad+N^2\lambda\left(2+(1+R)\lambda\right)\bar{\tau}_\theta^2\Big)+A^6\sigma_S^6\bar{\tau}_\theta^2\omega_X^6(\sigma_S^2+N\bar{\tau}_\theta^2)\Big]\Bigg)<0,
\end{aligned}
$$

and

$$\frac{1}{1+R}\left(\Delta^{NW}-\Delta^{PW}\right)=$$
$$\Bigg(RI^2N\bar{x}^2A^4\sigma_S^4\bar{\tau}_\theta^2\omega_X^2\left((\lambda I)^2N+A^2\sigma_S^2\omega_X^2\right)$$
$$\left[(\lambda I)^2N(\sigma_S^2+N\bar{\tau}_\theta^2)+A^2\sigma_S^2\omega_X^2(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)\right]^2\Bigg)\Bigg/$$
$$\Bigg(\Big[(1+R)\lambda^4I^6N^2(\sigma_S^2+N\bar{\tau}_\theta^2)^2+2(1+R)\lambda^2I^4NA^2\sigma_S^2\omega_X^2$$
$$\cdot(\sigma_S^2+N\bar{\tau}_\theta^2)(\sigma_S^2+\lambda N\bar{\tau}_\theta^2)+I^2A^4\sigma_S^4\omega_X^4\Big((1+R)\sigma_S^4$$
$$+\lambda N\left(2(1+R)+\lambda\right)\sigma_S^2\bar{\tau}_\theta^2+(1+R)\lambda^2N^2\bar{\tau}_\theta^4\Big)+A^6\sigma_S^8\bar{\tau}_\theta^2\omega_X^6\Big]$$
$$\cdot\Big[(1+R)\lambda^4I^6N^2(\sigma_S^2+N\bar{\tau}_\theta^2)^2+\lambda^2I^4NA^2\sigma_S^2\omega_X^2$$
$$\cdot(\sigma_S^2+N\bar{\tau}_\theta^2)^2\left(2(1+R)\sigma^2+N\left(1+2(1+R)\lambda\right)\bar{\tau}_\theta^2\right)$$
$$+I^2A^4\sigma_S^4\omega_X^4\Big((1+R)\sigma_S^4+N\left(1+\lambda\left(2(1+R)+\lambda\right)\right)\sigma_S^2\bar{\tau}_\theta^2$$
$$+\lambda N^2\left(2+(1+R)\lambda\right)\bar{\tau}_\theta^4\Big)+A^6\sigma_S^6\bar{\tau}_\theta^2\omega_X^6(\sigma_S^2+N\bar{\tau}_\theta^2)\Big]\Bigg)>0.$$

C.7 Negative externality: Proof of proposition 5.3

To show that any signal to news watchers inflicts a negative externality on price watchers, it suffices to look at condition (5.28) if it does not change sign for any N. I will prove that this condition always has a negative sign.

First note that Δ^{PW} (C.37) can only turn positive for sufficiently large $\bar{x}^2$, where the threshold value for $\bar{x}^2$ is given by (C.39). Thus, no other parameter of the model can make condition (5.28) positive. So, it will suffice to show that condition (5.28) is strictly negative for any (weakly positive) $\bar{x}^2$. Note that condition (5.28) is linear in $\bar{x}^2$. Defining $A\equiv\mathbb{E}_{pre}^{PW}\left[(\mu^{PW}-RP)/\tau^{PW}\right]^2/\bar{x}^2$, $B\equiv-\mathbb{V}_{pre}^{PW}\left((\mu^{PW}-RP)/\tau^{PW}\right)$, $D_1\equiv(\varepsilon_{\tau^2,N}^{PW}+\varepsilon_{\mathbb{E},N}^{PW})/N$, $D_2\equiv(\varepsilon_{\tau^2,N}^{PW}+\frac{1}{2}\varepsilon_{\mathbb{V},N}^{PW})/N$, $G^{-1}\equiv(\tau^{PW})^2/(1+R)$, $K^{-1}\equiv 1-B/(1+R)$, we can rewrite condition (5.28) as

$$AK(D_1+BD_2\tfrac{K}{1+R})\bar{x}^2+BD_2\tfrac{K^2}{1+R}(B+G). \tag{C.41}$$

We are interested in the signs of the two terms in (C.41). By their definition, $B<0$, $K>0$, $G<0$. In addition, $D_2<0$ by (C.27) and (C.29) (see third row in table 5.1, p. 165). Thus, the second term in (C.41) is strictly negative, $BD_2\frac{K^2}{1+R}(B+G)<0$. The other term is harder to evaluate, however, since $D_1<0$ by the sum of (C.27) and (C.29) (see third row in table 5.1 again). To find its sign, we can proceed in the following manner.

Setting the rewritten condition (C.41) equal to zero and solving out for $\bar{x}^2$, we find

$$(\bar{x}_0^{neg.ext.})^2 = -\frac{BD_2\frac{K^2}{1+R}(B+G)}{AK(D_1 + BD_2\frac{K}{1+R})}.$$

Thus, the sign of $AK(D_1 + BD_2\frac{K}{1+R})$ is the same as that of $(\bar{x}_0^{neg.ext.})^2$. By the equilibrium values (C.20) through (C.22) and (C.27) through (C.29) we find

$$\begin{aligned}
(\bar{x}_0^{neg.ext.})^2 \quad = \quad -\Bigg(& A^2\sigma_S^2\omega_X^4 \\
& \Big[\lambda^5 I^6 N^2(\sigma_S^2 + N\bar{\tau}_\theta^2)(\sigma_S^2 + 4N\bar{\tau}_\theta^2) + \lambda^3 I^4 N \\
& \quad \cdot A^2\sigma_S^2\omega_X^2\left(2\sigma_S^4 + N(11+\lambda)\sigma + S^2\bar{\tau}_\theta^2 + 2N^2(3+\lambda)\bar{\tau}_\theta^4\right) \\
& \quad +\lambda I^2 A^4\sigma_S^4\omega_X^4 \cdot \left(\sigma_S^4 + 3N(2+\lambda)\sigma_S^2\bar{\tau}_\theta^2 + 4\lambda N^2\bar{\tau}_\theta^4\right) \\
& \quad +2A^6\sigma_S^8\bar{\tau}_\theta^2\omega_X^6\Big] \\
& \Big[(1+R)\lambda^6 I^8 N^3(\sigma_S^2 + N\bar{\tau}_\theta^2)^3 + (1+R)\lambda^4 I^6 N^2 \\
& \quad \cdot A^2\sigma_S^2\omega_X^2(\sigma_S^2 + N\bar{\tau}_\theta^2)^2(3\sigma_S^2 + 2\lambda N\bar{\tau}_\theta^2) + \lambda^2 I^4 N \\
& \quad A^4\sigma_S^4\omega_X^4\Big(3(1+R)\sigma_S^6 + (1+R)N(3+4\lambda)\sigma_S^4\bar{\tau}_\theta^2 + \lambda N \\
& \qquad \cdot\sigma_S^2\bar{\tau}_\theta^2\left((1+R)N(4+\lambda) + \lambda\sigma_S^2\right) + (1+R)\lambda^2 N^3\bar{\tau}_\theta^2\Big) \\
& \quad +I^2 A^6\sigma_S^8\omega_X^6\Big((1+R)\sigma_S^4 + (1+R)N^2\lambda^2\bar{\tau}_\theta^2 \\
& \qquad +2\lambda N\sigma_S^2\bar{\tau}_\theta^2(1+R+\lambda\bar{\tau}_\theta^2)\Big) + A^8\sigma_S^1 2\bar{\tau}_\theta^4\omega_X^8\Big]\Bigg) \Bigg/ \\
\Bigg(& I^2\bar{\tau}_\theta^4(\lambda^2 I^2 N + A^2\sigma_S^2\omega_X^2) \\
& \left((\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2\right) \\
& \left((\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right) \\
& \Big[2(1+R)\lambda^7 I^8 N^4(\sigma_S^2 + N\bar{\tau}_\theta^2)^2 \\
& \quad +2(1+R)\lambda^5 I^6 N^3 A^2\sigma_S^2\omega_X^2(\sigma_S^2 + N\bar{\tau}_\theta^2) \\
& \qquad \left(4\sigma_S^2 + N(2+\lambda)\bar{\tau}_\theta^2\right) + \lambda^3 I^4 N^2 A^4\sigma_S^4\omega_X^4 \\
& \qquad \cdot\Big(\left(10 + (2-\lambda)\lambda + 2R(5+\lambda)\right)\sigma_S^4 + 2N\sigma_S^2\bar{\tau}_\theta^2 \\
& \qquad\quad \cdot\left(4 + (5-\lambda)\lambda + R(4+5\lambda)\right)\sigma_S^2\bar{\tau}_\theta^2 + 6(1+R)\lambda N^2\bar{\tau}_\theta^4\Big) \\
& \quad +2\lambda I^2 N A^6\sigma_S^6\omega_X^6\Big(\left(2 + (2-\lambda)\lambda + 2R(1+\lambda)\right)\sigma_S^4 \\
& \qquad +\lambda N\left(4 - \lambda + R(4+\lambda)\right)\sigma_S^2\bar{\tau}_\theta^2 + (1+R)\lambda^2 N^2\bar{\tau}_\theta^2\Big) \\
& +A^8\sigma_S^{10}\left((2(1+R)-\lambda)\sigma_S^2 + 2R\lambda N\bar{\tau}_\theta^2\right)\Big]\Bigg) \; < 0.
\end{aligned} \tag{C.42}$$

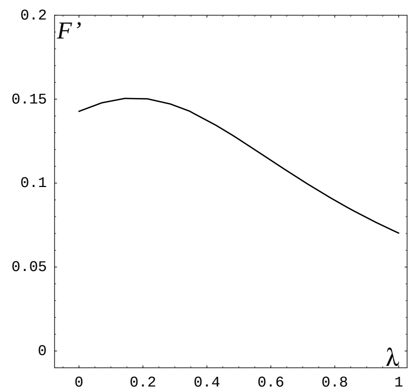


Figure C.1: Equilibria in Grossman and Stiglitz' (1980) model

So, this zero-point of condition (5.28) would lie in the strictly negative range of $\bar{x}^2$ if that existed. Therefore, $AK(D_1 + BD_2 \frac{K}{1+R}) < 0$ so that all terms in condition (5.28) are strictly negative, which concludes the proof. ■

C.8 Grossman and Stiglitz' (1980) version

We can compare the two-group equilibrium in section 5.4 to the equilibrium in Grossman and Stiglitz' (1980) model. Since Grossman and Stiglitz assume, too, that news watchers get perfect copies of the newspapers, their model remains a special case of the two-group model in section 5.4. There is a continuum of investors in Grossman and Stiglitz' world and investors may either watch the price or receive exactly one signal in addition. Thus, by setting $I = 1$ and $N = 1$ throughout the present model, Grossman and Stiglitz' version results. Accordingly, the cost of becoming a news watcher can be redefined as $F' = F + c$ here. Figure C.1 depicts the equilibrium share of news watchers $\lambda^*(F')$ as a function of the fixed information cost F'.[1] Just as in the present framework, multiple equilibria may arise in Grossman and Stiglitz' model, too. For high values of F, there are two possible equilibrium levels of λ. In addition, we can implement any equilibrium share of news watchers λ by varying F' in the example of figure C.1. This stands in contrast to proposition 5.4 which states that, as soon as news watchers can choose the number of newspapers, at least one investor must obtain less (or different) information in equilibrium. The reason for this difference is that the implicit equilibrium definition in Grossman and Stiglitz' variant of the model has suppressed the optimizing behavior of investors. It is merely a fixed point that equalizes *ex ante* utility. In

[1] Levels of F' on the vertical axis are expressed as shares of wealth. Parameter values are the same as in figure 5.2, except for $I = 1$, for $W = 1$, which takes the same value as in figure 5.3, and for F, which is endogenous here. See footnote 11 (p. 172).

the present framework, however, news watchers may have second thoughts once they have become news watchers. They can pay the fixed cost of joining the news watchers group, but then discover that their representative actually prefers to order zero news papers for everyone. This possibility for a second thought makes all the difference. As so often in game theoretically oriented models, the outcome depends on the structure of the game.

Apart from the no-equilibrium conjecture for fully revealing prices, discussed at large in the previous section 5.3, Grossman and Stiglitz (1980) introduced several further conjectures. Many of them concern the informativeness of the equilibrium price. Formally, the informativeness of a signal is its precision, which, in turn, is defined as the inverse of its *ex ante* variance. So, to investigate the informativeness of price, consider its variance

$$\begin{aligned}\mathbb{V}_{pre}(RP) &= \pi_S^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + \pi_X^2\omega_X^2 \\ &= \frac{\bar{\tau}_\theta^4\left[(\lambda I)^2 N + A^2\sigma_S^2\omega_X^2\right]^2\left[(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2\right]}{I^2\left[(\lambda I)^2 N(\sigma_S^2 + N\bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda N\bar{\tau}_\theta^2)\right]^2}.\end{aligned}$$

This follows from (5.16) in the text and lemma 5.3. For $N = I = 1$, the precision of RP is

$$\mathbb{V}_{pre}(RP)^{-1}\Big|_{N=I=1} = \frac{\left[\lambda^2(\sigma_S^2 + \bar{\tau}_\theta^2) + A^2\sigma_S^2\omega_X^2(\sigma_S^2 + \lambda\bar{\tau}_\theta^2)\right]^2}{\bar{\tau}_\theta^4\left[\lambda^2 + A^2\sigma_S^2\omega_X^2\right]^2\left[\lambda^2(\sigma_S^2 + \bar{\tau}_\theta^2) + A^2\sigma_S^4\omega_X^2\right]}. \tag{C.43}$$

It is not difficult to show that the precision of the price can be rising or falling in λ, and rising or falling in the precision of news watchers' signals σ_S^2—by taking the respective derivatives and playing with parameter values.[2] This result weakens Grossman and Stiglitz' (1980) conjectures 1 and 4 which asserted monotonous changes. The difference between the two models arises because Grossman and Stiglitz assume in their derivation that the realization of the equilibrium price depends on the expected value of the risky asset supply (see (A10) in Grossman and Stiglitz 1980), while the realization of equilibrium price in the present model is dependent on the realization of asset supply, and not its expectation (see (5.25)). The latter is consistent with our typical notion of a Walrasian equilibrium.

C.9 General model

Consider an economy with I investor who may acquire different amounts of information. Assume that the N^i signals which investor i purchases are independently drawn for every investor. That is, signals are not sold in perfect copies but in *individual* editions. Suppose that the equilibrium

[2] For example, use the values underlying figure 5.2 (as in footnote 11) to evaluate the according elasticities and then use $A = .1, \omega_X = 1, \lambda = .9$.

price takes a linear form of the type

$$RP = \pi_0 + \sum_{i=1}^{I} \pi_S^i \sum_{j=1}^{N^i} S_j^i - \pi_X X \tag{C.44}$$

for $I+2$ coefficients $\pi_0, \pi_S^1, ..., \pi_S^I, \pi_X$ to be determined. The vector of $\sum_i N^i$ signals is normally distributed where all signals S_j^i are conditionally independent given a realization of θ which is normally distributed itself. So, $(S_1^1, ..., S_{N^1}^1, ..., S_1^I, ..., S_{N^I}^I|\theta)^T \sim \mathcal{N}(\theta \cdot \iota_{\Sigma_i N^i}, \sigma_S^2 \cdot \mathbf{I}_{\Sigma_i N^i})$, where $\iota_{\Sigma_i N^i}$ is an $(\sum_i N^i) \times 1$ vector of ones and $\mathbf{I}_{\sum_i N^i}$ an identity matrix of dimension $(\sum_i N^i)$. Consider asset supply X to be independently normally distributed with $X \sim \mathcal{N}(\bar{x}, \omega_X^2)$.

A rational investor i departs from a joint normal distribution of all these random variables when updating her prior beliefs to posterior beliefs. Given any choice of N^i, investor i looks at the the joint distribution of the normally distributed random vector $(\theta; S_1, ..., S_{N^i}; RP)^T$. This vector has a prior mean

$$\bar{\mu}^i = \left(\bar{\mu}_\theta;\ \bar{\mu}_\theta, ..., \bar{\mu}_\theta;\ \pi_0 + \left(\sum_{k=1}^I \pi_S^k N^k\right) \bar{\mu}_\theta - \pi_X \bar{x}\right)^T$$

and a prior $(N^i+2) \times (N^i+2)$ variance-covariance matrix

$$\bar{\Sigma}^i = \begin{pmatrix} \bar{\tau}_\theta^2 & \bar{\tau}_\theta^2 \cdot \iota_{N^i}^T & \mathbb{C}\mathtt{ov}(RP, \theta) \\ \bar{\tau}_\theta^2 \cdot \iota_{N^i} & \mathbb{C}\mathtt{ov}(\mathbf{S}^i.\mathbf{S}^{i,T})_{N^i} & \mathbb{C}\mathtt{ov}^i(RP, S^i)\iota_{N^i} \\ \mathbb{C}\mathtt{ov}(RP, \theta) & \mathbb{C}\mathtt{ov}^i(RP, S^i)\iota_{N^i}^T & \mathbb{V}(RP) \end{pmatrix}.$$

$\mathbf{S}^i = (S_1^i, ..., S_N^i)^T$ is the vector of investor i's N^i signals, ι_{N^i} denotes an N^i vector of ones, while

$$\mathbb{C}\mathtt{ov}(\mathbf{S}^i.\mathbf{S}^{i,T})_{N^i} = \begin{pmatrix} \bar{\tau}_\theta^2 + \sigma_S^2 & \bar{\tau}_\theta^2 & \cdots & \bar{\tau}_\theta^2 \\ \bar{\tau}_\theta^2 & \bar{\tau}_\theta^2 + \sigma_S^2 & & \bar{\tau}_\theta^2 \\ \vdots & & \ddots & \vdots \\ \bar{\tau}_\theta^2 & \cdots & & \bar{\tau}_\theta^2 + \sigma_S^2 \end{pmatrix},$$

and

$$\begin{aligned} \mathbb{C}\mathtt{ov}(RP, \theta) &= \left(\sum_{k=1}^I \pi_S^k N^k\right) \bar{\tau}_\theta^2, \\ \mathbb{C}\mathtt{ov}^i(RP, S^i) &= \left(\sum_{k=1}^I \pi_S^k N^k\right) \bar{\tau}_\theta^2 + \pi_S^i \sigma_S^2, \\ \mathbb{V}(RP) &= \left(\sum_{k=1}^I \pi_S^k N^k\right)^2 \bar{\tau}_\theta^2 + \left(\sum_{k=1}^I (\pi_S^k)^2 N^k\right) \sigma_S^2 + \pi_X^2 \omega_X^2. \end{aligned}$$

Note that the vector of means only differs in its dimension across investors, whereas the variance-covariance matrix differs both in its dimension and its value due to an individual's π_S^i that comes into play through $\mathbb{C}\mathtt{ov}^i(RP, S^i)$.

As $\omega_X^2 \to 0$, the price becomes fully revealing. In this case, the variance of the price simplifies to $\mathbb{V}(RP) = (\sum_{k=1}^{I} \pi_S^k N^k)^2 \bar{\tau}_\theta^2 + (\sum_{k=1}^{I} (\pi_S^k)^2 N^k) \sigma_S^2$ since price contains no more exogenous noise. Suppose $\pi_S^k = \pi_S$ for all k. Then, applying fact C.1 yields posterior expectations

$$\mathbb{E}\left[\theta \,\middle|\, RP; s_1^i, ..., s_{N^i}^i\right] = \mu^i = \frac{\pi_S \bar{\mu}_\theta \sigma_S^2 + (RP - \pi_0 + \pi_X x)\bar{\tau}_\theta^2}{\pi_S(\sigma_S^2 + \bar{\tau}_\theta^2 \sum_{k=1}^{I} N^k)} \equiv \mu$$

and posterior variance

$$\mathbb{V}\left(\theta \mid RP; s_1^i, ..., s_{N^i}^i\right) = (\tau^i)^2 = \frac{\sigma_S^2 \bar{\tau}_\theta}{\sigma_S^2 + \bar{\tau}_\theta^2 \sum_{k=1}^{I} N^k} \equiv \tau^2.$$

So, posterior beliefs coincide for all investors. Applying this to price (5.9), we find $RP = \mu - \frac{A\bar{x}}{I}\tau^2$. Hence, $\pi_S = 1$ which proves the assertion $\pi_S^k = \pi_S \,\forall k$. Thus, the results are those of fact 5.1. This confirms the derivations in section 5.3.