RES Easter School: Behavioural Economics
Brasenose College Oxford, 22-25 March 2015

Reference-dependence and loss aversion II: Applications

Vincent P. Crawford
University of Oxford, All Souls College,
and University of California, San Diego

(with thanks to Juanjuan Meng)

UC San Diego

European Research Council
Established by the European Commission

Revised 8 March 2015

Reference-dependent preferences and cabdrivers' labor supply

The standard microeconomic model of labor supply is a special case of traditional consumer theory, with workers selling some of their endowment of leisure to "buy" income, at a price determined by the wage.

Reference-dependent preferences and cabdrivers' labor supply

The standard microeconomic model of labor supply is a special case of traditional consumer theory, with workers selling some of their endowment of leisure to "buy" income, at a price determined by the wage.

Some influential empirical work studies cabdrivers' labor supply.

Reference-dependent preferences and cabdrivers' labor supply

The standard microeconomic model of labor supply is a special case of traditional consumer theory, with workers selling some of their endowment of leisure to "buy" income, at a price determined by the wage.

Some influential empirical work studies cabdrivers' labor supply.

(Cabdrivers still choose their own hours, so their choices are closer to the standard model than the choices of most workers in modern economies.

They are likely to be approximately risk-neutral in their unpredictable daily earnings, so expected earnings per hour is a close analogue of the wage.)

Reference-dependent preferences and cabdrivers' labor supply

The standard microeconomic model of labor supply is a special case of traditional consumer theory, with workers selling some of their endowment of leisure to "buy" income, at a price determined by the wage.

Some influential empirical work studies cabdrivers' labor supply.

(Cabdrivers still choose their own hours, so their choices are closer to the standard model than the choices of most workers in modern economies.

They are likely to be approximately risk-neutral in their unpredictable daily earnings, so expected earnings per hour is a close analogue of the wage.)

A high "wage" early in a day suggests, treating early earnings as an indicator of earnings later that day, that you could earn more easily, by working longer that day.

Thus in the absence of implausibly large income effects, the neoclassical substitution effect predicts that, hours will have positive wage elasticity.

Even so, in the first such study, of New York City cabdrivers, Camerer et al. (1997 *QJE*) found a strongly negative elasticity of hours with respect to the "wage", with elasticities in subgroups between -0.503 and -0.269.

These negative elasticities reduce earnings, posing a puzzle for the model.

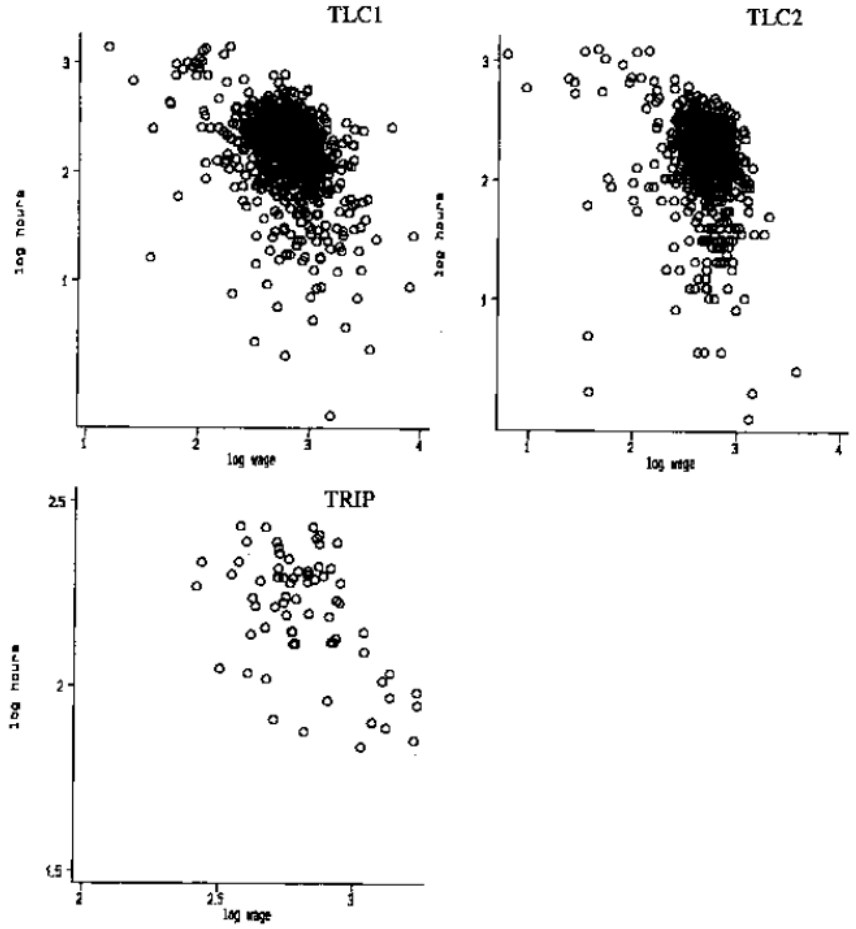

FIGURE I
Hours-Wage Relationships

They are neither a quirk of Camerer et al.’s dataset nor of cabdrivers; other studies yielded similar results, which appear to have similar explanations:

They are neither a quirk of Camerer et al.'s dataset nor of cabdrivers; other studies yielded similar results, which appear to have similar explanations:

- Kahneman, Knetsch, and Thaler's (1990 *JPE*) and other studies of the "endowment effect" and the "willingness-to-pay-willingness-to-accept gap".

- Odean's (1998 *Journal of Finance*) and Meng's (2013) studies of the "disposition effect": investors selling "winners" relative to purchase price but holding "losers", lowering their returns.

- Genesove and Mayer's (2001 *QJE*) study of the effects of house purchase prices (a sunk cost) on subsequent house asking prices.

- Farber's (2005 *JPE*, 2008 *AER*) studies of a new group of New York City cabdrivers, and other studies of cabdrivers around the world.

- Lien and Zheng's (2015 *AERP&*P) study of gamblers' stopping decisions.

E.g. Kahneman, Knetsch, and Thaler (1990 *JPE*) randomly gave mugs to half their subjects (“owners”) and nothing to the others (“nonowners”).

They then elicited selling and buying prices for owners or nonowners, using a procedure that gives subjects an incentive to reveal their true prices.

E.g. Kahneman, Knetsch, and Thaler (1990 *JPE*) randomly gave mugs to half their subjects (“owners”) and nothing to the others (“nonowners”).

They then elicited selling and buying prices for owners or nonowners, using a procedure that gives subjects an incentive to reveal their true prices.

With random assignment and symmetric information about the mugs, in a large enough sample the standard neoclassical supply and demand curves should be mirror images of each other.

E.g. Kahneman, Knetsch, and Thaler (1990 *JPE*) randomly gave mugs to half their subjects ("owners") and nothing to the others ("nonowners").

They then elicited selling and buying prices for owners or nonowners, using a procedure that gives subjects an incentive to reveal their true prices.

With random assignment and symmetric information about the mugs, in a large enough sample the standard neoclassical supply and demand curves should be mirror images of each other.

But average buying price of nonowners was about $3.50 and average selling price of owners was about $7.00: Way too big to be random.

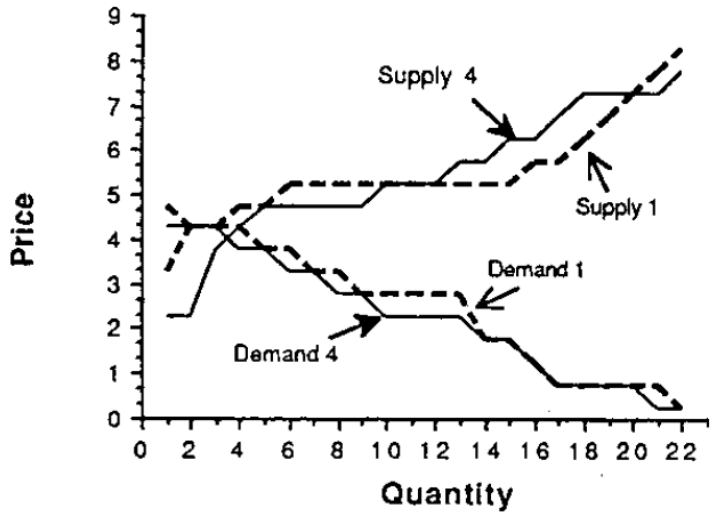

FIG. 1.—Supply and demand curves, markets 1 and 4

This result has been replicated many times, with the gap almost always large and in the same direction.

(But see the “maybe-no-gap” findings of Plott and Zeiler, 2005, 2007 *AER* and the 2011 *AER* Comment and Reply by Isoni et al. and Plott and Zeiler.)

Such a “willingness-to-pay/willingness-to-accept” gap is inconsistent with choices based on preferences only over *levels* of consumption of mugs and money (with the latter viewed as a proxy for all other consumption).

But the gap is gracefully explained by a reference-dependent model with reference points for mugs and money defined by expectations.

In such a model loss aversion predicts not only the existence of a gap, but also its almost always observed positive direction.

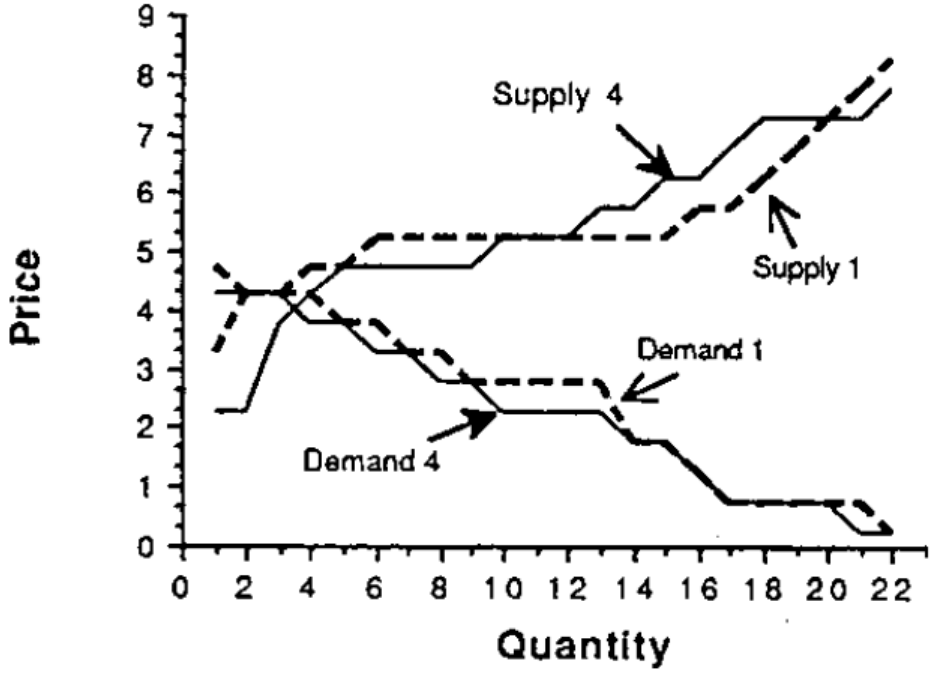

FIG. 1.—Supply and demand curves, markets 1 and 4

E.g. Farber (2005 *JPE*), inspired by Camerer et al.'s study, collected and analyzed data on a new set of New York City cabdrivers.

As is normal in the labor literature, Farber imposed driver homogeneity, with individual variation in determinants of drivers' choices but not preferences.

E.g. Farber (2005 *JPE*), inspired by Camerer et al.'s study, collected and analyzed data on a new set of New York City cabdrivers.

As is normal in the labor literature, Farber imposed driver homogeneity, with individual variation in determinants of drivers' choices but not preferences.

His main findings were:

- Before controlling for driver fixed effects, the probability of stopping work is significantly related to income realized on a given day, but

- Driver fixed effects and other relevant controls render this effect statistically insignificant, and

- The probability of stopping is significantly related to cumulative hours.

Farber (2008 *AER*) estimated a structural model with explicit daily income targets, again imposing homogeneity of drivers' preferences.

He treated the (not directly observable) targets as latent variables with driver-specific means and driver-independent variance.

He assumed, for computational reasons, that both mean and variance of income are constant across days of the week, thus allowing the target to vary across days for a given driver, but only as a random effect.

(This assumption is strongly rejected in the data, with Thursdays' through Sundays' mean incomes systematically higher than those of other days.

Farber included day-of-the-week dummies in his main specifications of the stopping probability equation, but this turns out to be an imperfect substitute for allowing the income target to vary across days of the week.)

Farber (2008 *AER*) found that a sufficiently rich parameterization of his income-targeting model has a better fit than a standard neoclassical model.

The estimated probability of stopping increases significantly and substantially once the income target is reached.

But the estimated model cannot reconcile the strong increase in stopping probability at the target with the aggregate smoothness of the relationship between stopping probability and realized income.

Further, the random effects in drivers' targets have very high estimated variances, from which he concluded that income targets are too unstable and imprecisely estimated to yield a useful reference-dependent model.

As Farber (2008 *AER*) suggests, a finding that labor supply is reference-dependent would have significant policy implications:

"Evaluation of much government policy regarding tax and transfer programs depends on having reliable estimates of the sensitivity of labor supply to wage rates and income levels. To the extent that individuals' levels of labor supply are the result of optimization with reference-dependent preferences, the usual estimates of wage and income elasticities are likely to be misleading."

As Farber (2008 *AER*) suggests, a finding that labor supply is reference-dependent would have significant policy implications:

"Evaluation of much government policy regarding tax and transfer programs depends on having reliable estimates of the sensitivity of labor supply to wage rates and income levels. To the extent that individuals' levels of labor supply are the result of optimization with reference-dependent preferences, the usual estimates of wage and income elasticities are likely to be misleading."

If reference-dependence is really there, even if its effects wash out on average, ignoring it (as has been the custom in labor economics) may yield biased estimates even of the models' "neoclassical" coefficients that measure the effects of permanent, predictable parameter changes.

But despite a number of empirical studies, the literature has not converged on the extent to which the evidence supports reference-dependence.

To explain the negative elasticities, Camerer et al. proposed that drivers have daily income targets and work until they reach their target for the day.

To explain the negative elasticities, Camerer et al. proposed that drivers have daily income targets and work until they reach their target for the day.

They therefore work less on days with high wages:

- Literal income targeting yields a wage elasticity of hours of -1.

- With less exact targeting the elasticity remains < 1, but moves toward 0.

Income targeting is in the spirit of Kahneman and Tversky's (1979 *Econometrica*, 1991 *QJE*) theory of reference-dependent preferences.

They suggested that people care not about levels of income or consumption as in standard consumer theory, but about how it changes.

They also suggested that such "reference-dependence" is accompanied by "loss aversion", whereby a decrease in consumption below the reference point (a "loss") hurts more than an equal increase above it (a "gain") helps.

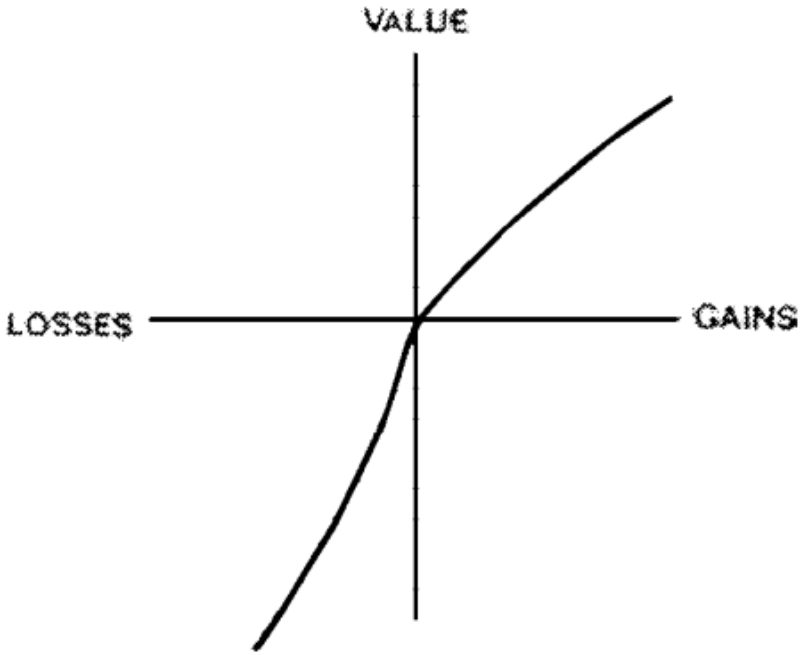

FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

They are less important here, but Kahneman and Tversky also suggested, based on the psychology of perception, that reference-dependent preferences reflect “diminishing sensitivity”, with the graph concave for gains but convex for losses; and “nonlinear probability weighting” (not shown).

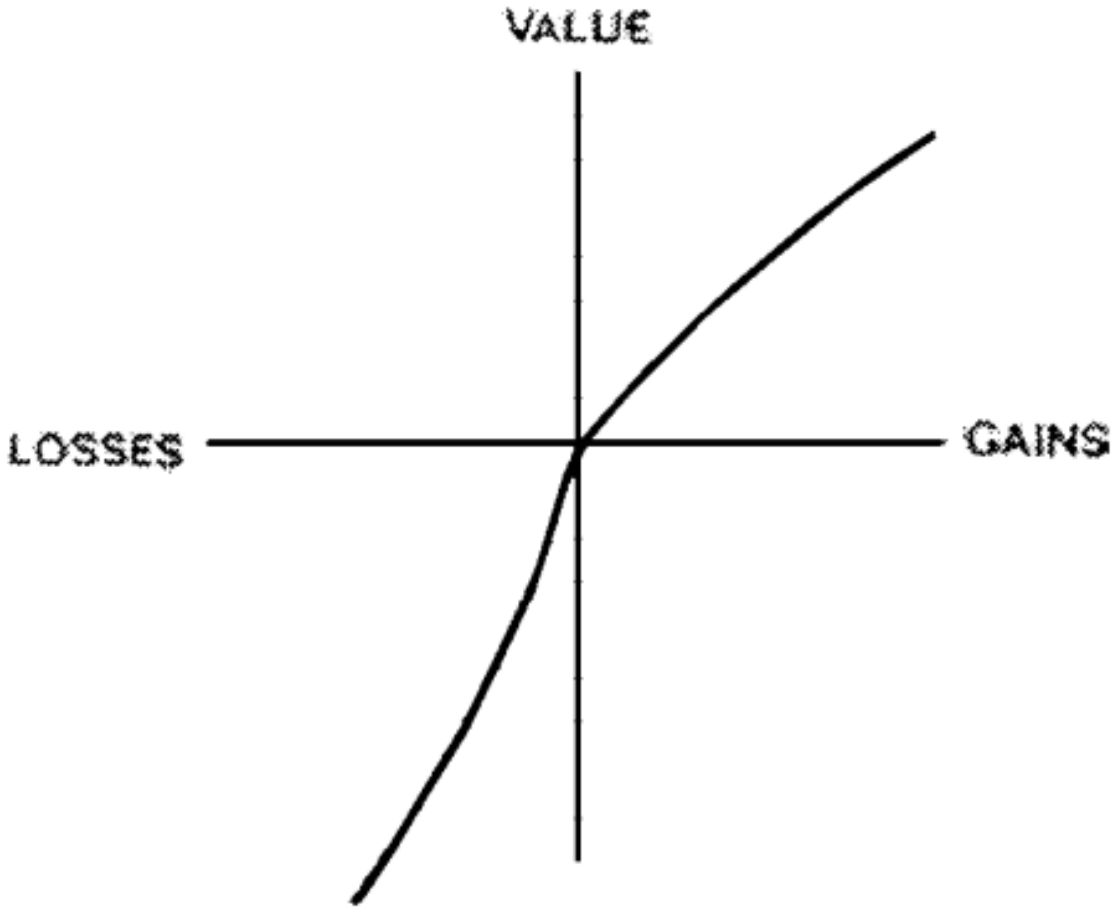

FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

Even though the negative elasticities that reference-dependence and loss aversion entail reduce earnings, reference-dependence is not irrational per se: It simply alters what people are rational about, from levels to changes.

Even though the negative elasticities that reference-dependence and loss aversion entail reduce earnings, reference-dependence is not irrational per se: It simply alters what people are rational about, from levels to changes.

Kahneman and Tversky (1979 *Econometrica*) stress that prospect theory's emphasis on changes is a basic aspect of human nature:

> An essential feature of the present theory is that the carriers of value are changes in wealth or welfare, rather than final states. This assumption is compatible with basic principles of perception and judgment. Our perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitudes. When we respond to attributes such as brightness, loudness, or temperature, the past and present context of experience defines an adaptation level, or reference point, and stimuli are perceived in relation to this reference point (Helson (1964)). Thus, an object at a given temperature may be experienced as hot or cold to the touch depending on the temperature to which one has adapted. The same principle applies to non-sensory attributes such as health, prestige, and wealth. The same level of wealth, for example, may imply abject poverty for one person and great riches for another depending on their current assets.

Kahneman (2003 *AER*) gives powerful visual examples of the perceptual importance of changes.

FIGURE 7. AN ILLUSION OF ATTRIBUTE SUBSTITUTION

Kahneman (2003 *AER*) gives powerful visual examples of the perceptual importance of changes.

FIGURE 7. AN ILLUSION OF ATTRIBUTE SUBSTITUTION

(The two “horsies” are the same size; and crucially, the illusion persists even when we understand how it was created.)

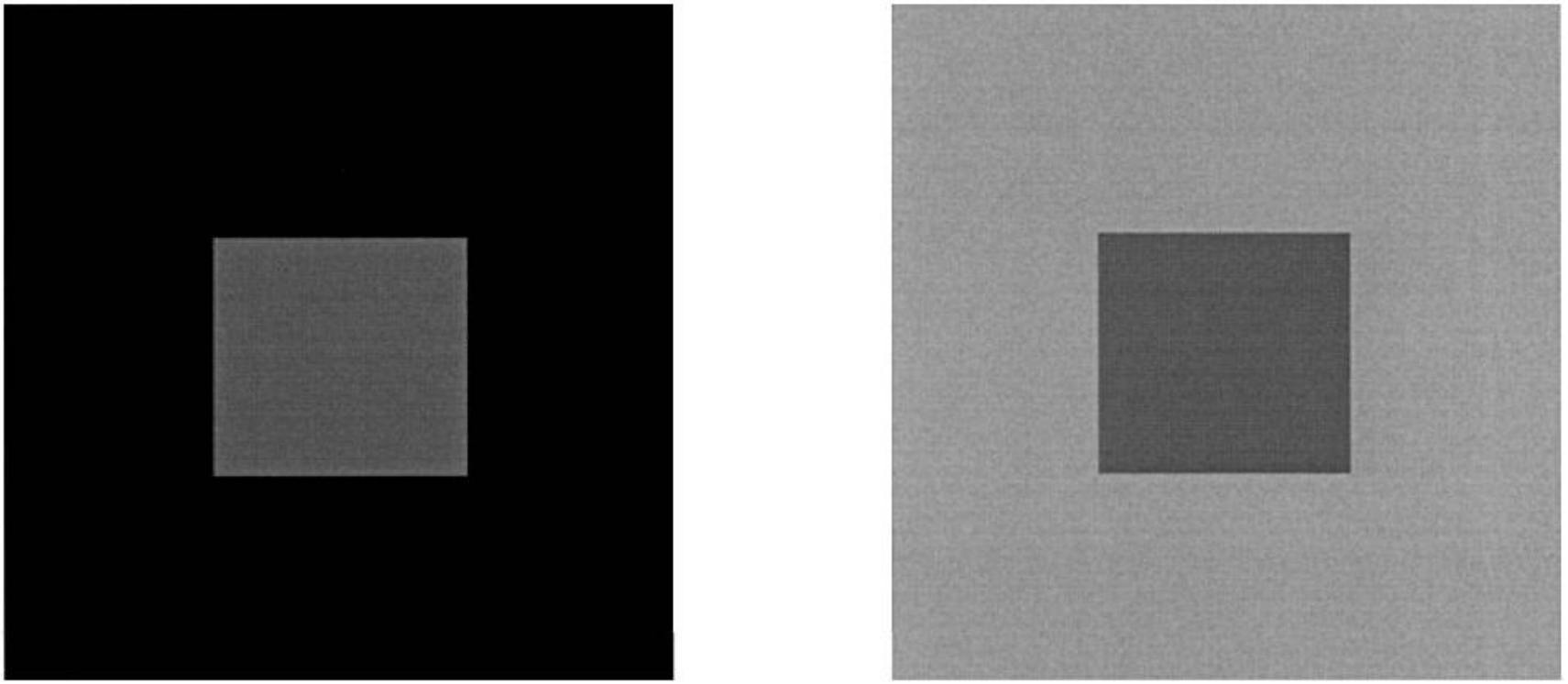

FIGURE 5. REFERENCE-DEPENDENCE IN THE PERCEPTION OF BRIGHTNESS

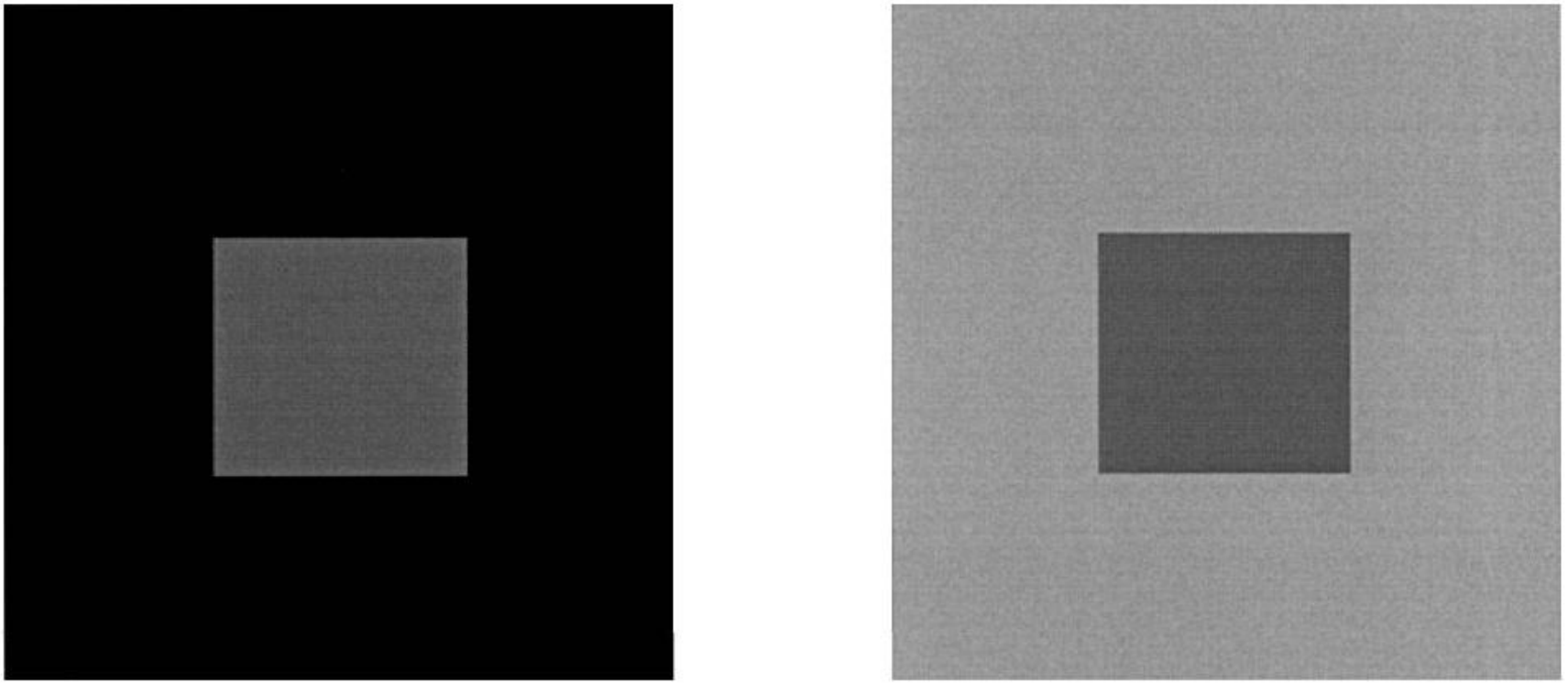

FIGURE 5. REFERENCE-DEPENDENCE IN THE PERCEPTION OF BRIGHTNESS

(The two inner squares are equally bright; and again the illusion persists.)

Reference-dependence is a promising candidate to explain the above empirical anomalies, which are inconsistent with standard consumer theory.

But there are several issues to resolve:

- In the field, even people who care about changes in consumption also care about its level.

Reference-dependence is a promising candidate to explain the above empirical anomalies, which are inconsistent with standard consumer theory.

But there are several issues to resolve:

- In the field, even people who care about changes in consumption also care about its level.

- Kahneman and Tversky's 1979 model was one-dimensional (generalized in 1991), but consumer theory is used in two or more dimensions.

Reference-dependence is a promising candidate to explain the above empirical anomalies, which are inconsistent with standard consumer theory.

But there are several issues to resolve:

- In the field, even people who care about changes in consumption also care about its level.

- Kahneman and Tversky's 1979 model was one-dimensional (generalized in 1991), but consumer theory is used in two or more dimensions.

- Gains and losses must necessarily be measured relative to some reference point, and a useful model requires a definite specification.

Reference-dependence is a promising candidate to explain the above empirical anomalies, which are inconsistent with standard consumer theory.

But there are several issues to resolve:

- In the field, even people who care about changes in consumption also care about its level.

- Kahneman and Tversky's 1979 model was one-dimensional (generalized in 1991), but consumer theory is used in two or more dimensions.

- Gains and losses must necessarily be measured relative to some reference point, and a useful model requires a definite specification.

- Despite the observed negative elasticities, few economists believe that a permanent, predictable increase in the wage would decrease hours; and studies like Fehr and Goette's (2007 *AER*) tend to confirm that.

Kőszegi and Rabin (2006 *QJE*) addressed these issues by proposing a reference-dependent model designed for economic applications.

In their theory, as applied to cabdrivers' labor supply:

- A driver's preferences reflect both the standard consumption utility of income and leisure and reference-dependent "gain-loss" utility, with their relative importance tuned by an estimated parameter.

Kőszegi and Rabin (2006 *QJE*) addressed these issues by proposing a reference-dependent model designed for economic applications.

In their theory, as applied to cabdrivers' labor supply:

- A driver's preferences reflect both the standard consumption utility of income and leisure and reference-dependent "gain-loss" utility, with their relative importance tuned by an estimated parameter.

- A driver has a daily target for hours as well as income, and he is loss-averse in both dimensions, with working longer than the hours target a loss, just as earning less than the income target is.

Kőszegi and Rabin (2006 *QJE*) addressed these issues by proposing a reference-dependent model designed for economic applications.

In their theory, as applied to cabdrivers' labor supply:

- A driver's preferences reflect both the standard consumption utility of income and leisure and reference-dependent "gain-loss" utility, with their relative importance tuned by an estimated parameter.

- A driver has a daily target for hours as well as income, and he is loss-averse in both dimensions, with working longer than the hours target a loss, just as earning less than the income target is.

- Most importantly, the targets are endogenized by setting them equal to a driver's theoretical rational expectations of hours and income, in Kőszegi and Rabin's notion of "preferred personal equilibrium".

Kőszegi and Rabin (2006 *QJE*) addressed these issues by proposing a reference-dependent model designed for economic applications.

In their theory, as applied to cabdrivers' labor supply:

- A driver's preferences reflect both the standard consumption utility of income and leisure and reference-dependent "gain-loss" utility, with their relative importance tuned by an estimated parameter.

- A driver has a daily target for hours as well as income, and he is loss-averse in both dimensions, with working longer than the hours target a loss, just as earning less than the income target is.

- Most importantly, the targets are endogenized by setting them equal to a driver's theoretical rational expectations of hours and income, in Kőszegi and Rabin's notion of "preferred personal equilibrium".

Abeler et al. (2011 *AER*) conducted a careful experimental test of Kőszegi and Rabin's expectational view of reference points, and largely confirmed it.

Crawford and Meng (2011 *AER*) estimated a model based on Kőszegi and Rabin's theory, using Farber's (2005 *JPE*, 2008 *AER*) data, again imposing homogeneity of drivers' preferences.

Crawford and Meng followed Kőszegi and Rabin in allowing consumption as well as gain-loss utility, and hours as well as income targets, with similar separability and functional structure assumptions.

Crawford and Meng (2011 *AER*) estimated a model based on Kőszegi and Rabin's theory, using Farber's (2005 *JPE*, 2008 *AER*) data, again imposing homogeneity of drivers' preferences.

Crawford and Meng followed Kőszegi and Rabin in allowing consumption as well as gain-loss utility, and hours as well as income targets, with similar separability and functional structure assumptions.

Crawford and Meng followed Kőszegi and Rabin in dropping diminishing sensitivity and nonlinear probability weighting, to focus on loss aversion.

Crawford and Meng (2011 *AER*) estimated a model based on Kőszegi and Rabin's theory, using Farber's (2005 *JPE*, 2008 *AER*) data, again imposing homogeneity of drivers' preferences.

Crawford and Meng followed Kőszegi and Rabin in allowing consumption as well as gain-loss utility, and hours as well as income targets, with similar separability and functional structure assumptions.

Crawford and Meng followed Kőszegi and Rabin in dropping diminishing sensitivity and nonlinear probability weighting, to focus on loss aversion.

They deviated from Kőszegi and Rabin's theory mainly in assuming, for simplicity, that the targets are point expectations rather than distributions.

(Kőszegi and Rabin needed target distributions for expectations-based reference-dependence to have an effect, in their model without errors.)

Crawford and Meng (2011 *AER*) estimated a model based on Kőszegi and Rabin's theory, using Farber's (2005 *JPE*, 2008 *AER*) data, again imposing homogeneity of drivers' preferences.

Crawford and Meng followed Kőszegi and Rabin in allowing consumption as well as gain-loss utility, and hours as well as income targets, with similar separability and functional structure assumptions.

Crawford and Meng followed Kőszegi and Rabin in dropping diminishing sensitivity and nonlinear probability weighting, to focus on loss aversion.

They deviated from Kőszegi and Rabin's theory mainly in assuming, for simplicity, that the targets are point expectations rather than distributions.

(Kőszegi and Rabin needed target distributions for expectations-based reference-dependence to have an effect, in their model without errors.)

Crawford and Meng's model, like Kőszegi and Rabin's, is consistent with rationality, with a concave objective function.

Its only important deviation from a neoclassical model is adding changes in income and leisure to their levels in the domain of preferences.

Crawford and Meng generally followed Farber's (2005 *JPE*, 2008 *AER*) econometric strategies, except that they followed Kőszegi and Rabin in treating the targets as rational expectations, not as latent variables.

Crawford and Meng operationalized drivers' unobservable expectations by finding natural sample proxies for them, with *limited* endogeneity problems.

Crawford and Meng generally followed Farber's (2005 *JPE*, 2008 *AER*) econometric strategies, except that they followed Kőszegi and Rabin in treating the targets as rational expectations, not as latent variables.

Crawford and Meng operationalized drivers' unobservable expectations by finding natural sample proxies for them, with *limited* endogeneity problems.

Closing the model this way avoids Farber's computational problems, which were what led him to conclude that his income targets are too unstable and imprecisely estimated to yield a useful reference-dependent model.

The additional structure from treating the targets as rational expectations also allows tests of the model by looking for systematic shifts in drivers' stopping decisions associated with the proxied targets.

Details

Treating each day separately as in all previous analyses, consider the preferences of a given driver during his shift on a given day.

I and *H* denote income earned and hours worked that day.

I^r and H^r denote income and hours targets for the day.

Total utility, $V(I, H|I^r, H^r)$, is a weighted average of consumption utility $U_1(I)$ + $U_2(H)$ and gain-loss utility $R(I, H|I^r, H^r)$, with weights $1 - \eta$ and η ($0 \le \eta \le 1$):

$$(1)\quad V(I,H \mid I^r, H^r) = (1-\eta)(U_1(I) + U_2(H)) + \eta R(I,H \mid I^r, H^r)\ ,$$

Details

Treating each day separately as in all previous analyses, consider the preferences of a given driver during his shift on a given day.

I and H denote income earned and hours worked that day.

I^r and H^r denote income and hours targets for the day.

Total utility, $V(I, H|I^r, H^r)$, is a weighted average of consumption utility $U_1(I)$ + $U_2(H)$ and gain-loss utility $R(I, H|I^r, H^r)$, with weights $1 - \eta$ and η ($0 \le \eta \le 1$):

$$(1)\quad V(I,H \mid I^r, H^r) = (1-\eta)(U_1(I) + U_2(H)) + \eta R(I,H \mid I^r, H^r),$$

where gain-loss utility

$$(2)\quad R(I,H \mid I^r, H^r) = 1_{(I-I^r \le 0)}\lambda(U_1(I) - U_1(I^r)) + 1_{(I-I^r > 0)}(U_1(I) - U_1(I^r)$$
$$+ 1_{(H-H^r \ge 0)}\lambda(U_2(H) - U_2(H^r)) + 1_{(H-H^r < 0)}(U_2(H) - U_2(H^r)).$$

(1)-(2) incorporate several of Kőszegi and Rabin's provisional assumptions:

- Consumption utility is additively separable across income and hours, with $U_1(\cdot)$ increasing in I, $U_2(\cdot)$ decreasing in H, and both concave.

(1)-(2) incorporate several of Kőszegi and Rabin's provisional assumptions:

- Consumption utility is additively separable across income and hours, with $U_1(\cdot)$ increasing in I, $U_2(\cdot)$ decreasing in H, and both concave.

- Gain-loss utility is also additively separable, determined component by component by differences in realized and target consumption utilities.

  Thus when the realized wage equals its expected value, so realized income and hours equal their target values, the model trivially reduces to a neoclassical model.

(1)-(2) incorporate several of Kőszegi and Rabin's provisional assumptions:

- Consumption utility is additively separable across income and hours, with $U_1(\cdot)$ increasing in I, $U_2(\cdot)$ decreasing in H, and both concave.

- Gain-loss utility is also additively separable, determined component by component by differences in realized and target consumption utilities.

  Thus when the realized wage equals its expected value, so realized income and hours equal their target values, the model trivially reduces to a neoclassical model.

- Gain-loss utility is a linear function of those utility differences, ruling out Prospect Theory's "diminishing sensitivity" as in a leading case Kőszegi and Rabin sometimes focus on (their Assumption A3').

(1)-(2) incorporate several of Kőszegi and Rabin's provisional assumptions:

- Consumption utility is additively separable across income and hours, with $U_1(\cdot)$ increasing in I, $U_2(\cdot)$ decreasing in H, and both concave.

- Gain-loss utility is also additively separable, determined component by component by differences in realized and target consumption utilities.

  Thus when the realized wage equals its expected value, so realized income and hours equal their target values, the model trivially reduces to a neoclassical model.

- Gain-loss utility is a linear function of those utility differences, ruling out Prospect Theory's "diminishing sensitivity" as in a leading case Kőszegi and Rabin sometimes focus on (their Assumption A3').

- Losses have a constant weight λ relative to gains, "the coefficient of loss aversion," which is the same for income and hours. Empirically, $\lambda \approx 2$ to 3.

The model allows a simple characterization of a driver's optimal stopping decision with targets for hours as well as income.

The optimal stopping decision of a driver who expects the wage to remain constant at w^e maximizes reference-dependent utility $V(I, H|I^r, H^r)$ as in (1) and (2), subject to the linear menu of income-hours combinations $I = w^e H$.

(When $U_1(\cdot)$ and $U_2(\cdot)$ are concave, $V(I, H|I^r, H^r)$ is concave in I and H for any given targets I^r and H^r (this depends on ruling out "diminishing sensitivity").)

The model allows a simple characterization of a driver's optimal stopping decision with targets for hours as well as income.

The optimal stopping decision of a driver who expects the wage to remain constant at w^e maximizes reference-dependent utility $V(I, H|I^r, H^r)$ as in (1) and (2), subject to the linear menu of income-hours combinations $I = w^e H$.

(When $U_1(\cdot)$ and $U_2(\cdot)$ are concave, $V(I, H|I^r, H^r)$ is concave in I and H for any given targets I^r and H^r (this depends on ruling out "diminishing sensitivity").)

The optimal stopping decision is characterized by a first-order condition, generalized to allow kinks at the reference points.

The driver starts with hours and income both 0, and continues working as long as his expected wage w^e exceeds the marginal rate of substitution.

Table 1 lists the marginal rates of substitution in the four possible gain-loss domains, expressed as hours disutility costs of an additional unit of income.

(On boundaries, marginal rates of substitution are replaced by generalized derivatives whose left- and right-hand values equal the interior values.)

These results all depend on the separability and functional structure assumptions in Kőszegi and Rabin's and Crawford and Meng's models, but they have sensible (though messier) generalizations.

TABLE 1—MARGINAL RATES OF SUBSTITUTION WITH REFERENCE-DEPENDENT PREFERENCES BY DOMAIN

	Hours gain $(H < H^r)$	Hours loss $(H > H^r)$
Income gain $(I > I^r)$	$-U_2'(H)/U_1'(I)$	$-[U_2'(H)/U_1'(I)][1-\eta+\eta\lambda]$
Income loss $(I < I^r)$	$-[U_2'(H)/U_1'(I)]/[1-\eta+\eta\lambda]$	$-U_2'(H)/U_1'(I)$

TABLE 1—MARGINAL RATES OF SUBSTITUTION WITH REFERENCE-DEPENDENT PREFERENCES BY DOMAIN

	Hours gain ($H < H^r$)	Hours loss ($H > H^r$)
Income gain ($I > I^r$)	$-U_2'(H)/U_1'(I)$	$-[U_2'(H)/U_1'(I)][1 - \eta + \eta\lambda]$
Income loss ($I < I^r$)	$-[U_2'(H)/U_1'(I)]/[1 - \eta + \eta\lambda]$	$-U_2'(H)/U_1'(I)$

When $\eta \approx 0$ or $\lambda = 1$, as when the realized wage equals its expected value, the model trivially reduces to a neoclassical model.

TABLE 1—MARGINAL RATES OF SUBSTITUTION WITH REFERENCE-DEPENDENT PREFERENCES BY DOMAIN

	Hours gain ($H < H^r$)	Hours loss ($H > H^r$)
Income gain ($I > I^r$)	$-U_2'(H)/U_1'(I)$	$-[U_2'(H)/U_1'(I)][1-\eta+\eta\lambda]$
Income loss ($I < I^r$)	$-[U_2'(H)/U_1'(I)]/[1-\eta+\eta\lambda]$	$-U_2'(H)/U_1'(I)$

When $\eta \approx 0$ or $\lambda = 1$, as when the realized wage equals its expected value, the model trivially reduces to a neoclassical model.

Even when $\eta > 0$, $\lambda > 1$, and the wage deviates from the expected wage, when hours and income are both in the gains or losses domain the marginal rate of substitution is the same as for consumption utilities alone.

TABLE 1—MARGINAL RATES OF SUBSTITUTION WITH REFERENCE-DEPENDENT PREFERENCES BY DOMAIN

	Hours gain $(H < H^r)$	Hours loss $(H > H^r)$
Income gain $(I > I^r)$	$-U_2'(H)/U_1'(I)$	$-[U_2'(H)/U_1'(I)][1-\eta+\eta\lambda]$
Income loss $(I < I^r)$	$-[U_2'(H)/U_1'(I)]/[1-\eta+\eta\lambda]$	$-U_2'(H)/U_1'(I)$

When $\eta \approx 0$ or $\lambda = 1$, as when the realized wage equals its expected value, the model trivially reduces to a neoclassical model.

Even when $\eta > 0$, $\lambda > 1$, and the wage deviates from the expected wage, when hours and income are both in the gains or losses domain the marginal rate of substitution is the same as for consumption utilities alone.

But when $\eta > 0$, $\lambda > 1$, and hours and income are in opposite domains, the marginal rate of substitution equals the consumption-utility trade-off times either $(1 - \eta + \eta\lambda)$ (> 1 when $\lambda > 1$) or $1/(1 - \eta + \eta\lambda)$.

(The tradeoff favors work more than the neoclassical tradeoff in the income loss/hours gain domain, but less in the hours loss/income gain domain.)

In Figure 1 the driver starts in the lower right-hand corner with hours and income 0, and continues as long as w^e > the marginal rate of substitution.

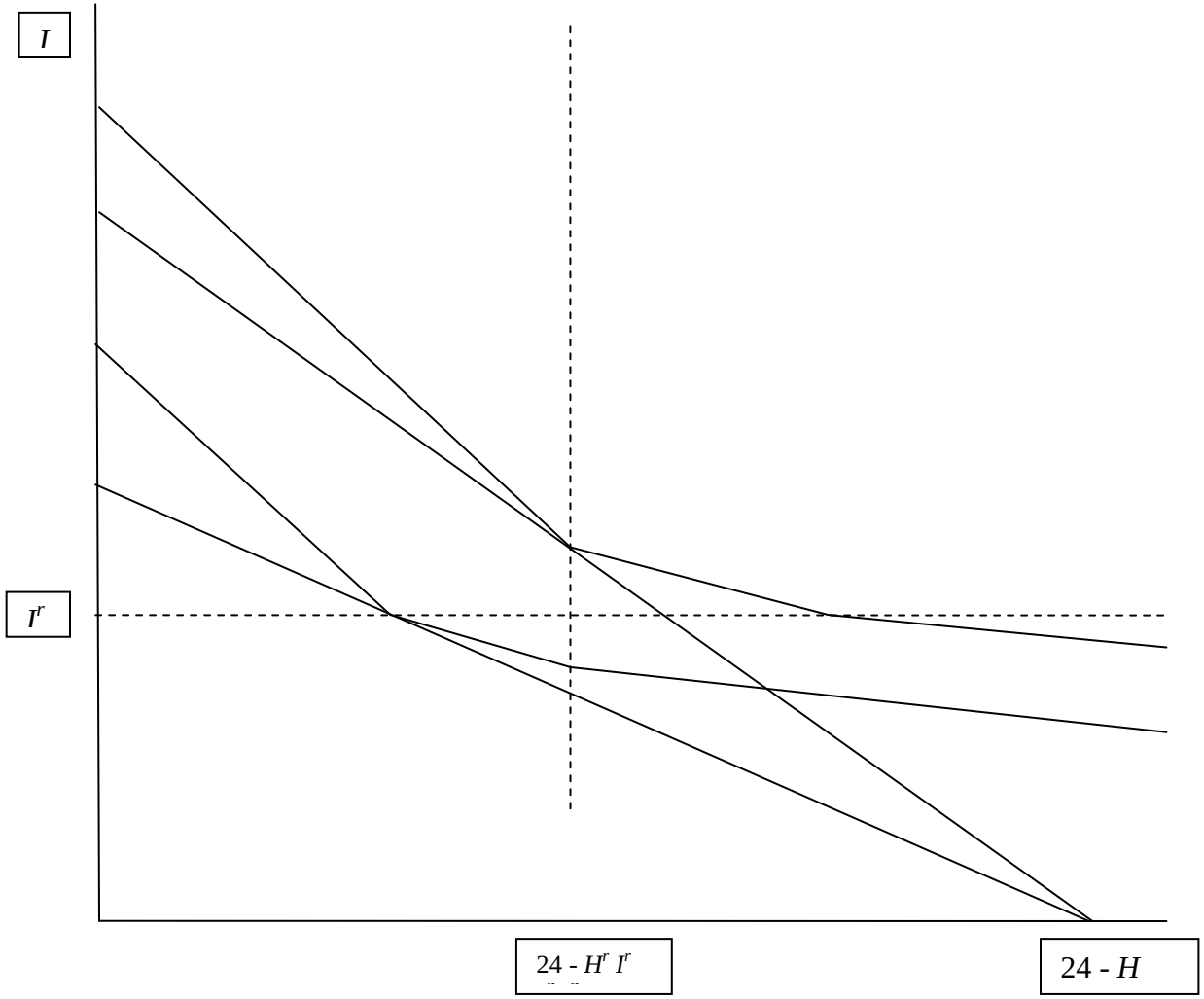

Figure 1: A Reference-dependent Driver's Stopping Decision

The driver anticipates passing through a series of domains such that the work-leisure tradeoff favors work less and less as hours and income accumulate, reflecting the concavity of reference-dependent utility in *I* and *H*.

The probability of stopping is more strongly influenced by hours or income, depending on which target is reached first.

The driver anticipates passing through a series of domains such that the work-leisure tradeoff favors work less and less as hours and income accumulate, reflecting the concavity of reference-dependent utility in *I* and *H*.

The probability of stopping is more strongly influenced by hours or income, depending on which target is reached first.

Crawford and Meng's structural estimates suggest that most drivers stop near the second target they reach on a given day (could go either way):

- Thus most drivers stop at the hours target on a good day (wage higher than expected), with wage elasticity 0.

- But most drivers stop at the income target on a bad day, with wage elasticity -1: even though they value income and are rational in the generalized, reference-dependent sense of the term.

The driver anticipates passing through a series of domains such that the work-leisure tradeoff favors work less and less as hours and income accumulate, reflecting the concavity of reference-dependent utility in *I* and *H*.

The probability of stopping is more strongly influenced by hours or income, depending on which target is reached first.

Crawford and Meng's structural estimates suggest that most drivers stop near the second target they reach on a given day (could go either way):

- Thus most drivers stop at the hours target on a good day (wage higher than expected), with wage elasticity 0.

- But most drivers stop at the income target on a bad day, with wage elasticity -1: even though they value income and are rational in the generalized, reference-dependent sense of the term.

With rational-expectations targets and roughly symmetric distributions, the average wage elasticity in the aggregate sample is roughly -½ (compare Camerer et al.'s estimates of -0.503 and -0.269).

Crawford and Meng's model yields estimates of the targets that are stable and sufficiently precise to yield a useful reference-dependent model.

Crawford and Meng's model yields estimates of the targets that are stable and sufficiently precise to yield a useful reference-dependent model.

Their estimated model shows that Kőszegi and Rabin's rational-expectations targets and distinction between consumption and gain-loss utility can gracefully reconcile:

- The negative wage elasticity of hours found by Camerer et al. (1997 *QJE*) and Farber (2005 *JPE*, 2008 *AER*).

Crawford and Meng's model yields estimates of the targets that are stable and sufficiently precise to yield a useful reference-dependent model.

Their estimated model shows that Kőszegi and Rabin's rational-expectations targets and distinction between consumption and gain-loss utility can gracefully reconcile:

- The negative wage elasticity of hours found by Camerer et al. (1997 *QJE*) and Farber (2005 *JPE*, 2008 *AER*).

- The positive relationship between *expected* wage, earnings, and hours found by Fehr and Goette (2007 *AER*) and others.

Crawford and Meng's model yields estimates of the targets that are stable and sufficiently precise to yield a useful reference-dependent model.

Their estimated model shows that Kőszegi and Rabin's rational-expectations targets and distinction between consumption and gain-loss utility can gracefully reconcile:

- The negative wage elasticity of hours found by Camerer et al. (1997 *QJE*) and Farber (2005 *JPE*, 2008 *AER*).

- The positive relationship between *expected* wage, earnings, and hours found by Fehr and Goette (2007 *AER*) and others.

- The smoothness of the relationship between stopping probability and realized income Farber found, via the heterogeneity of realized wages.

Supplement: A reference-dependent driver's labor supply curve

Figure 2 compares the labor-supply curves for a neoclassical driver and a reference-dependent driver with the same consumption utility functions.

The solid curve is the neoclassical supply curve, while the dashed curve is the reference-dependent one.

The shape of the curve depends on which target has a larger influence on the stopping decision, which depends in turn on the relationship between the neoclassical optimal stopping point (that is, the stopping point that maximizes consumption utility alone) and the targets.

Figure 2 illustrates the case suggested by Crawford and Meng's estimates:

For wages that reconcile the income and hours targets as at point D, the neoclassically optimal income and hours are higher than the targets, so the driver stops at his second-reached target.

Whenever the wage is to the left of point D, the hours target is reached before the income target, and vice versa.

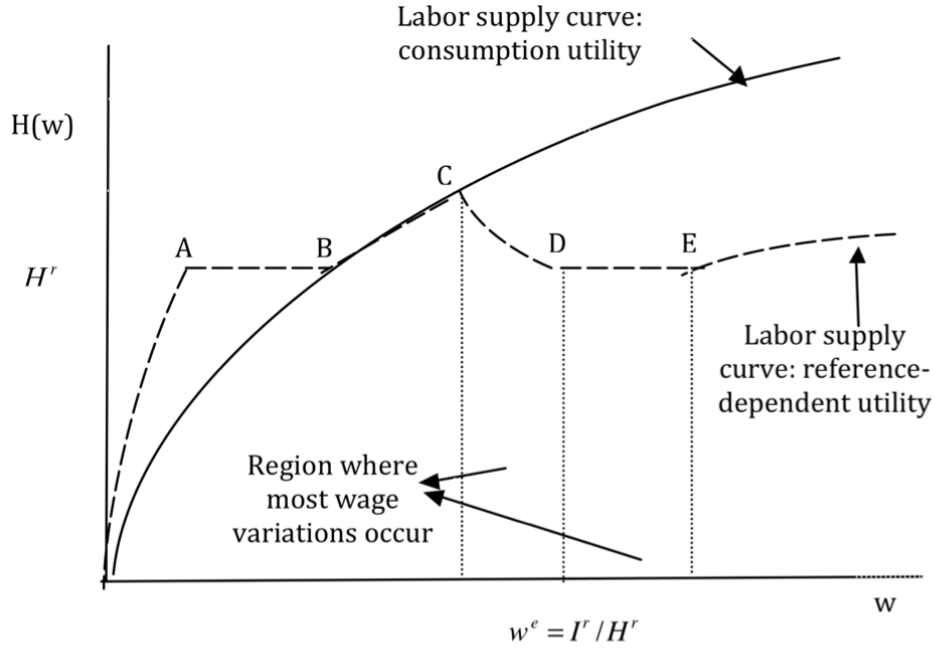

Figure 2: A Reference-dependent Driver's Labor Supply Curve

As Figure 2 illustrates, reference-dependent labor supply is non-monotonic.

When the wage is to the left of point A, the higher cost of income losses raises the incentive to work above its neoclassical level.

Along segment AB labor supply is determined by the kink at the hours target, which is reached first.

Along segment BC the neoclassical optimal stopping point is above the hours but below the income target, so the gain-loss effects cancel out, and reference-dependent and neoclassical labor supply coincide.

Along segment CD labor supply is determined by the kink at the income target, which is reached second, so the wage elasticity of hours is negative.

Along segment DE labor supply is determined by the kink at the hours target, which is reached second.

Finally, when the wage is to the right of point E, the higher cost of hours losses lowers the incentive to work below its neoclassical level.

Most realized wages fall close to point D.