Family Welfare Cultures

Gordon B. Dahl* Andreas Ravndal Kostol† Magne Mogstad‡

July 12, 2013

Abstract: Strong intergenerational correlations in various types of welfare use have fueled a long standing debate over whether welfare dependency in one generation causes welfare dependency in the next generation. Some claim a culture has developed in which welfare use reinforces itself through the family, because parents on welfare provide information about the program to their children, reduce the stigma of participation, or invest differentially in child development. Others argue the determinants of poverty or poor health are correlated across generations, so that children's welfare participation is associated with, but not caused by, parental welfare use. However, there is little empirical evidence to sort out these claims. In this paper, we investigate the existence and importance of family welfare cultures in the context of Norway's disability insurance (DI) system. To overcome the challenge of correlated unobservables across generations, we take advantage of random assignment of judges to DI applicants whose cases are initially denied. Some appeal judges are systematically more lenient, which leads to random variation in the probability a parent will be allowed DI. Using this exogenous variation, we find strong evidence that welfare use in one generation causes welfare use in the next generation: when a parent is allowed DI, their adult child's participation over the next five years increases by 6 percentage points. This effect grows over time, rising to 12 percentage points after ten years. Using our estimates, we simulate the total reduction in DI participation from a policy which makes the screening process more stringent; the intergenerational link amplifies the direct effect on parents at the margin of program entry, leading to long-run participation rates and program costs which are substantially lower than would otherwise be expected. The detailed nature of our data allows us to explore the mechanisms behind the causal intergenerational relationship; we find suggestive evidence against stigma and parental investments and in favor of children learning from a parent's experience with the DI program.

Keywords: Intergenerational welfare transmission, welfare cultures, disability insurance
JEL codes: I38, J62, H53

*Department of Economics, University of California San Diego; Research Department, Statistics Norway. E-mail: gdahl@ucsd.edu

†Research Department, Statistics Norway; Department of Economics, University of Bergen. E-mail: kos@ssb.no

‡Department of Economics, University College London; Research Department, Statistics Norway. E-mail: magne.mogstad@gmail.com

1 Introduction

Strong intergenerational correlations in various types of welfare use have fueled a long standing debate over whether welfare dependency in one generation causes welfare dependency in the next generation. Some policymakers and researchers have argued that a culture has developed in which welfare use reinforces itself through the family.[1] There are at least three pathways which could drive a culture of welfare within families: parents on welfare may supply information about the program to their children, reduce the stigma of participation, or invest differentially in child development. Each of these pathways imply that it is the parent's experience with welfare programs that creates a intergenerational link. An alternative explanation is that the determinants of poverty or poor health are correlated across generations in ways which have nothing to do with a welfare culture, but which nonetheless translate into similar participation rates within families. This explanation says that while a child's use of welfare may be correlated with a parent's use, it is not caused by the parent's welfare participation.

Estimating whether welfare dependency in one generation causes welfare dependency in the next generation has proven difficult given the likelihood of correlated unobservables across generations.[2] On top of this, it is often difficult to access large datasets on welfare use which link family members together across generations. These empirical challenges have meant that existing research has largely focused on intergenerational correlations in various types of welfare use. Black and Devereux (2011), in their Handbook of Labor Economics chapter, summarize the state of the literature well: "while the intergenerational correlations in welfare receipt are clear, there is much less evidence that a causal relationship exists."

In this paper, we investigate the existence and importance of family welfare cultures, where the take up of a welfare program by one generation causes increased participation in the next generation. We exploit a policy which randomizes the probability that parents receive welfare in combination with a unique source of population panel data. We estimate the causal relationship in welfare participation across generations in the context of Norway's disability insurance (DI) system. Our focus on DI receipt is highly policy relevant, as it is now one of the largest transfer programs in most industrialized countries. In the U.S., for example,

[1] For example, in his 1992 State of the Union Address, President George Bush said "Welfare was never meant to be a lifestyle; it was never meant to be a habit; it was never supposed to be passed from generation to generation like a legacy."

[2] Researchers have documented strong intergenerational patterns for a variety of socioeconomic variables (see e.g. Black and Devereux, 2011, Lee and Solon, 2009, Mazumder, 2005, Oreopoulos, Page, and Stevens, 2006), highlighting the difficulty of separating out correlations within families from causal effects. Bjorklund, Lindahl, and Plug (2006) show that both pre- and postbirth factors contribute substantially to intergenerational transmissions of socioeconomic variables. Levine and Zimmerman (1996) show a large portion of the observed correlation in AFDC participation can be explained by intergenerational correlations in income and other family characteristics. Pepper (2000) illustrates the difficulty in drawing causal inferences about intergenerational welfare transmission from observational data.

outlays for DI exceed those for food stamps, traditional cash welfare, or the EITC.[3] For families without small children, DI is often the only cash benefit available after unemployment benefits run out and it has therefore become an increasingly important component of the social safety net. Over the past 50 years, DI rolls have steadily risen from less than 1% to over 5% of the adult population in the U.S., from 1% to 7% in the U.K, and from 2% to almost 10% in Norway. Many have argued these increases are fiscally unsustainable, especially as current DI recipients are younger and have longer life expectancies on average compared to previous cohorts of recipients (e.g., Autor and Duggan, 2006; Burkhauser and Daly, 2012).

The key to our research design is that the DI system in Norway randomly assigns judges to DI applicants whose cases are initially denied. Some appeal judges are systematically more lenient, which leads to random variation in the probability an individual will be allowed DI. We utilize this exogenous variation to see if the DI participation of parents affects the probability their adult children subsequently apply for and are awarded DI. Our approach takes advantage of the fact that appeal judges are randomly assigned and therefore their leniency in a parent's case is unrelated to any other intergenerational factors (such as poverty or health) which might influence the DI participation of their children. To assess the internal validity of our research design, we perform a number of robustness checks, all of which suggest the identifying assumptions of independence, exclusion and monotonicity hold.

As our measure of judge leniency, we use the average allowance rate in the other cases a judge has handled. This leniency measure is highly predictive of the judge's decision in the current case, but as we document, uncorrelated with observable case characteristics. Using this random variation as an instrument, we find that welfare dependency in one generation causes welfare dependency in the next generation. When a parent is allowed DI because of a lenient judge, their adult child's participation rate increases by 6 percentage points over the next five years. This intergenerational welfare transmission amplifies over time; the effect of parental DI participation on their adult child's participation rate reaches 12 percentage points ten years after the judge's decision. By comparison, we calculate only one percent of these children would have been on DI if their parents had been denied DI. Consistent with this increase in adult children's welfare dependency, we find that parental DI receipt decreases the probability that a child will work or pursue higher education.

Our findings have important implications for the evaluation of welfare reforms, as any changes will affect not only the current generation, but their children as well. We use our estimates to simulate the

[3] In 2011 the U.S. paid out \$129 billion to 10.6 million disabled workers and their families, with an additional \$33 billion worth of disability benefits from the SSI program for poor Americans and \$90 billion in Medicaid for disabled workers (OASDI Trustees Report, 2012). By way of comparison, in the U.S. in 2011 the cash assistance portion of TANF paid out \$10 billion to 4.6 million participants, SNAP (food stamps) paid out \$80 billion to 46.5 million participants and the EITC paid out \$62 billion to 27 million working families. In 2009, DI payments constituted 1.8% of GDP in the U.S. and 2.3% of GDP across the European OECD-countries (OECD, 2010).

total reduction in DI participation from a policy which makes the screening process more stringent. In the early years after a tightening of the screening process, most of the reduction in DI participation can be attributed to the direct effect on parents, as there is little opportunity for children to learn and respond to their parent's DI experience. In contrast, the intergenerational effect grows over time; after ten years, the increase in children's participation accounts for almost half of the total reduction in DI rolls. In terms of program expenditure, it is important to capture this intergenerational effect, since few individuals exit DI after entering and the children are much younger than their parents when they enter DI.

We further use the rich Norwegian data to explore possible mechanisms. Since our baseline sample consists of children who are age-eligible for DI (at least 18 years old), our estimates cannot be attributed to differential parental investments during childhood. Our results are also not driven by differential investments as young adults, since the intergenerational relationship remains strong even when we exclude children who live at home or focus on children who are least 25 years of age. When we look at an alternative, smaller sample of children who are under 18 at the time of their parent's appeal decision, we still find parental DI participation substantially increases the probability that children will subsequently apply for and be awarded DI. Taken together, these findings suggest that differential investments by parents on welfare is not a key reason for the existence of family welfare cultures, at least in the context of DI.

We also find two pieces of evidence against the hypothesis that our findings are due to a drop in social stigma resulting from parental DI use. First, other forms of stigmatized welfare use by a child do not change after a parent is allowed DI, in contrast to what a model of general social stigma would imply. Second, the estimated effect of a parent's DI experience on a child's DI participation increases over time. However, many parents who were initially denied re-apply and are eventually allowed DI, which would suggest the gap in stigma between the treatment group (initially allowed parents) and the control group (initially denied parents) should shrink over time. In contrast, a model in which children learn from their parent's cumulative experience fits these time patterns. In such a model, children first learn about the initial rejection of a parent's DI application, and then as initially denied parents begin to re-apply for DI, their children additionally learn the process is time consuming, risky and increasingly costly.

Our paper complements a growing literature on the causes and consequences of the growth in DI rolls (for a review, see Autor and Duggan, 2006, Autor, 2011). To date, research has largely focused on estimating the work capacity and labor supply elasticity of DI recipients.[4] Yet despite a recent surge in research on

[4]See e.g. Autor and Duggan (2003), Borghans, Gielen, and Luttmer (2012), Bound (1989), Campolieti and Riddell (2012), French and Song (2013), Gruber (2000), Kostøl and Mogstad (2013), Maestas, Mullen, and Strand (2013), Parsons (1991), Moore (2011), von Wachter, Song, and Manchester (2011).

this topic, less is known about what causes individuals to apply for DI, why disability rolls have risen so dramatically, and how the receipt of DI affects individuals on margins other than labor force participation.[5] Our study provides some of the first causal evidence on what influences DI applications and what the effects of DI participation are for children of recipients. The magnitude of our estimates suggest that intergenerational transmission could play an important role in explaining the dramatic rise in DI participation over the past few decades.

Our study is also related to a small set of papers that have used assignment of judges or examiners in different contexts. Two studies using U.S. data and a similar research design have looked at how DI receipt affects labor supply.[6] Maestas, Mullen, and Strand (2013) use variation in the leniency of initial examiners in the U.S. and find that DI receipt substantially reduces earnings and employment of applicants. Exploiting the leniency of appeal judges in the U.S., French and Song (2013) find comparable labor supply effects of DI receipt among appellants. When applying this research design to the Norwegian data, our labor supply effects are quite similar to those found in the U.S., which indicates that the counterfactual labor outcomes for parents are comparable across the two countries. What makes our study unique is the ability to link the judicial decisions to a wide range of variables for both parents and their children. This allows us to provide novel evidence on whether and how welfare use in one generation causes welfare use in the next generation.

The remainder of the paper proceeds as follows. Section 2 discusses the challenges in estimating intergenerational welfare transmission and our experimental research design. In Section 3, we describe our data, provide institutional background, and compare the DI program in Norway with that of the U.S. Section 4 presents our main findings on intergenerational welfare transmission and reports robustness checks. Section 5 presents a policy simulation and Section 6 explores possible mechanisms. The final section offers some concluding remarks.

2 Identifying Intergenerational Welfare Transmission

2.1 Threats to Identification and Previous Research

In the spirit of Bertrand, Luttmer, and Mullainathan (2000), our definition of a family welfare culture is that take up of a welfare program by one generation causes increased participation in the next generation.

[5]Autor and Duggan (2006) discuss a number of possible explanations for the rise in DI rolls. There also exists a small body of evidence on entry responses to changes in DI benefits, wages, or local labor market conditions, including Black, Daniel, and Sanders (2002), Bratberg (1999), Campolieti (2004), Gruber (2000), and Rege, Telle, and Votruba (2009). None of these studies consider the role played by intergenerational welfare transmission.

[6]Assignment of judges or examiners has also been used in other contexts, such as the effect of incarceration on employment and earnings (Kling, 2006) and the effect of foster care placement on delinquency and crime (Doyle, 2007, 2008).

This can be modeled by relating child i's latent demand (and latent qualification) for a social program, P_i^{c*}, to their parent's actual participation P_i^p:

$$P_i^{c*} = \alpha^c + \beta^c P_i^p + \delta^c x_i^c + \varepsilon_i^c \tag{1}$$

where the superscripts c and p denote child and parent variables and coefficients. A child participates in the welfare program if $P_i^{c*} > 0$. In addition to the parent's decision, a child's participation also depends on a variety of other observable (x_i^c) and unobservable (ε_i^c) variables, such as demographic characteristics, parental characteristics, and the child's earnings capacity, health, and attitudes.

Of course, a similar equation can be written for the parent's social program decision:

$$P_i^{p*} = \alpha^p + \beta^p P_i^g + \delta^p x_i^p + \varepsilon_i^p \tag{2}$$

where the new superscript g denotes child i's grandparent. Some of the observed x_i^p variables could also directly affect P_i^{c*} and would therefore be included in x_i^c.

A bias in the family welfare culture parameter, β^c, can arise due to unobserved factors which are correlated across generations. This becomes apparent when substituting a parent's choice resulting from equation (2) into equation (1):

$$P^{c*} = \alpha^c + \beta^c I(\alpha^p + \beta^p P_i^g + \delta^p x_i^p + \varepsilon_i^p > 0) + \delta^c x_i^c + \varepsilon_i^c. \tag{3}$$

where $I(\cdot)$ is the indicator function. This formulation makes clear that if $corr(\varepsilon_i^p, \varepsilon_i^c | x_i^c, x_i^p) \neq 0$, there will be a bias. For example, low earnings potential could be correlated across generations due to unobservable factors common to the parent and child, such as bad neighborhoods or low quality schools. As another example, since there is a genetic component to health, certain physical ailments could reduce work capacity within families in ways unrelated to program participation. These correlations in unobservables could incorrectly lead a researcher to believe there is a family welfare culture, when in fact the patterns are simply due to intergenerational correlations in adverse environments or poor health.

This same reasoning extends to prior generations as well. Because equation (3) is recursive, it includes a variable for the participation of a child's grandparent, which itself depends on the participation of prior generations and a vector of observable (x_i^g) and unobservable (ε_i^g) variables. If $corr(\varepsilon_i^g, \varepsilon_i^c | x_i^c, x_i^p, x_i^g) \neq 0$ this can additionally bias the family welfare culture parameter. The potential for this type of bias is suggested by studies which document multi-generational correlations in a variety of variables such as income, poverty, education, and occupation (Black and Devereux, 2011, Lee and Solon, 2009). There is also evidence on multi-generational links in health status due to shared genes; the genetic expression of some of these conditions even skip a generation (for a review, see Bird, 2007).

Because many factors associated with welfare use are likely to be correlated across generations, the data demands for OLS estimation of equation (1) to yield causal evidence are high. One needs to have an exhaustive set of child and parent characteristics, as well as relevant controls for both sets of grandparents (and potentially prior generations as well). In Table 1, we show the intergenerational correlation in DI use and its sensitivity to the inclusion of controls. We use cross sectional data from Norway in 2008, restricting our attention to parents and adult children who are both age-eligible for DI. Column 1 finds that a child's DI participation more than doubles, rising by 3.6 percentage points, if a parent has participated in the DI program. Column 2 adds in the prior generation as well, and finds a small but statistically significant effect of any grandparent's DI use on the child above and beyond the effect from their parent. The final column adds in control variables for a variety of child, parent, and grandparent characteristics. These controls cut the estimated coefficient on parental DI use by almost a third and illustrate the sensitivity of OLS estimates to omitted variable bias.

Table 1: OLS Estimates of Intergenerational Welfare Transmission.

	Child DI use (P_i^c)		
	(1)	(2)	(3)
Parent DI use (P_i^p)	0.036***	0.035***	0.025***
	(0.001)	(0.001)	(0.001)
Grandparent DI use (P_i^g)		0.005***	0.004***
		(0.000)	(0.000)
Additional controls?	NO	NO	YES
N	1,022,507	1,022,507	1,022,507
Dependent mean	0.03	0.03	0.03

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the family level.

Notes: Data come from a cross section covering all Norwegian residents in 2008. Sample restricted to children age 23 and older with parents age 60 or younger (and a grandparent who was alive during the period 1967-2010); these age restrictions mirror those for the baseline estimation sample five years after the parent's appeal decision (see Section 3). DI use in each generation defined to be equal to 1 if the individual is currently receiving DI benefits (except for grandparents, which is defined as having ever received DI benefits). The third column controls flexibly for child, parent and grandparent characteristics (age, gender, education, foreign born, marital status, earnings history, and municipality fixed effects).

A number of studies have used observational data to estimate models like equation (1).[7] As we do, they find strong intergenerational correlations. While these studies have helped researchers and policymakers

[7] See e.g. Duncan, Hill, and Hoffman (1988), Solon, Corcoran, Gordon, and Laren (1988), Moffitt (1992), Antel (1992), Page and Stevens (2002) and Page (2004).

better describe intergenerational patterns in various types of welfare use, a causal interpretation remains elusive. As is well understood, such regressions cannot distinguish state dependence (the causal effect of program participation) from that of unobserved heterogeneity (correlated unobservables across generations).

There have been a few attempts to find instruments for parental welfare use (such as state benefit levels or local labor market conditions), include family fixed effects, or impose structural restrictions to estimate the causal intergenerational link.[8] Pepper (2000) illustrates the difficultly in drawing credible inferences from observational data. Using a nonparametric bounds analysis, he shows that without prior information about the selection problem, the data are not informative about intergenerational welfare use. Even imposing strong assumptions or using standard instruments, he finds the bounds are wide and the point estimates are noisy and often inconsistent across specifications.

2.2 Experimental Setting and Research Design

In this subsection, we begin by reviewing key facts regarding the DI program in Norway. We then provide institutional details and empirical evidence on the disability determination process, documenting in particular that the system generates random variation in DI awards. We further describe how we will use this exogenous variation to estimate the intergenerational link in DI participation.

The Norwegian DI Program

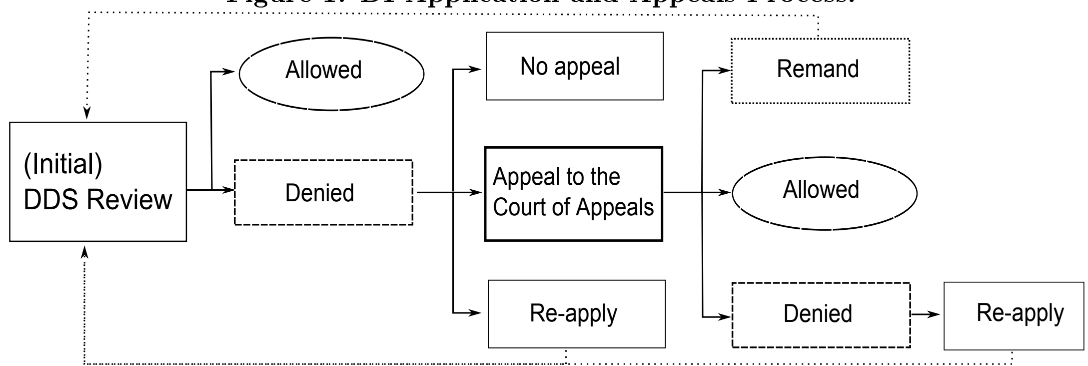
In Norway, DI benefits are designed to provide partial earnings replacements to all workers under the full retirement age who are unable to engage in substantial gainful activity because of a medically determined physical or mental impairment that has lasted for at least a year. The DI program is part of the broader Social Security System and is financed through employer- and employee-paid taxes. The level of DI benefits received is determined using a formula based on an individual's earnings history. The proportion of income that is replaced decreases as past earnings increase so that low-wage workers replace a larger fraction of their earnings than do high-wage workers.

The disability determination process is a multi-step process. Figure 1 shows the different steps. The first step is the submission of an initial application to the Social Security Administration office for the Disability Determination Stage (DDS) review. If the applicant meets the non-medical criteria (such as age and prior employment requirements), disability examiners and medical staff assess written medical evidence regarding the applicant's ability to perform work-related activities. Examiners take into account health

[8] See Levine and Zimmerman (1996), Gottschalk (1996), Pepper (2000), Beaulieu, Duclos, Fortin, and Rouleau (2005) and Bratberg, Nilsen, and Vaage (2012).

status, age, education, and work experience as well as the transferability of the applicant's skills. If the disability examiner concludes that the applicant cannot be expected to engage in any substantial gainful activity, a disability award is made. Approximately 75% of claims are awarded at this first step. Cases that are more difficult to judge (such as mental illness and low back pain) are often denied at this step.

Figure 1: DI Application and Appeals Process.



If the DI claim is initially denied, the individual may appeal the decision within 2 months to the Court of Appeals. About 25% of all denials are appealed. DI appeals are reviewed by Administrative Law Judges (ALJs). The ALJ must consider the application using the same criteria as the initial determination, but the applicant may present new information in writing. Judges can either allow a case, deny a case, or issue a remand (which means the case is sent back to the DDS Review stage to be re-evaluated with updated information).[9] Approximately 15% of all claims that were appealed are allowed at the ALJ level. If the case is denied at the ALJ level, the applicant can always choose to start a new DI case by re-applying to the DDS Review stage.[10]

Random Assignment of DI Cases to Judges

In Norway, the hearing of appeals is centralized in Oslo, where cases are handled for the entire country. Prior to 1998, there was only one department. Afterwards, there were four equally-sized departments; however, there is no specialization in the four departments and all judges are housed in the same building. Within each department, the assignment of a case to an Administrative Law Judge is done by the department head without knowing the content of the case, as stipulated in the rules set forth for the Administrative Law Court

[9] Remands are uncommon, accounting for only 5 percent of appeal outcomes. In our baseline analysis, we code remanded cases as rejected. In a robustness check, we code remanded cases as allowed or denied based on their eventual outcome after they are reconsidered by the DDS case worker with updated information and the results are similar.

[10] If a case is denied at the ALJ level, it can also be appealed to the higher courts, but very few individuals exercise this option.

since its inception in 1967. The rules state that assignment should be done "by the drawing of lots." In practice, cases are assigned on a rotating basis depending on the date they are received and the alphabetical ordering of a judge's last name.[11]

Our setting has several attractive features: (i) the handling of cases is centralized in one location, (ii) judges do not specialize by medical condition, region of country, or other aspects of the case, (iii) the judge assesses the written evidence on the appellant's case; there is never any personal contact between the judge and those who appeal, and (iv) an individual cannot choose an alternate judge after being assigned a judge.

The key to our design is not only that the assignment of judges is random, but also that some judges are more lenient than others. We measure judge leniency based on the average allowance rate in all other cases a judge has handled.[12] To construct the judge leniency measure, we calculate the leave-out mean judge allowance rate and regress this measure on fully interacted time and department dummies; this is because the randomization occurs among the pool of judges within each department. We use the residual from this regression as our judge leniency measure. This approach controls for any differences over time or across departments in the quality of applicants and the leniency of the judges.

Verifying Random Assignment

Table 2 empirically verifies that the hearing office complied with the random allocation procedure. This table conducts the same type of statistical test that would be done for an actual experiment to verify compliance with randomization. We find strong empirical support for the claim that the DI system in Norway randomly assigns judges to individuals who appeal their cases. The first column documents that demographic, work and health variables are highly predictive of whether an appealed case will be allowed. Column 3 examines whether our measure of judge leniency can be predicted by these same characteristics. Even though the set of characteristics are highly predictive of case outcomes, they are not statistically related to the leniency of the judge assigned to a case: none of the 14 variables are statistically significant at the 5% significance level and the variables are not jointly significant either. In fact, the point estimates are close to zero, and taken together, the variables explain only 0.24 percent of the variation in our measure of judge leniency.[13]

A natural question is why some judges are more lenient than others. While we do not have detailed characteristics of the judges, we do know the number of cases they have handled. Appendix Figure A.1 plots

[11] We verified these rules with the current Head of the Administrative Law Court, Knut Brofoss. The rules are explained in "Veileder for Saksbehandlingen i Trygderetten" (Guidelines for Processing Cases in the Court of Appeals).

[12] Throughout the paper, we calculate the leniency measure based on all the cases a judge has handled, and not just those cases appearing in our estimation sample. On average, judges handle 380 cases.

[13] The coefficient on age, while close to zero, is statistically significant at the 10% level. Given the number of covariates we consider, this is not surprising, since the probability of observing one p-value at this level by chance alone is large.

Table 2: Testing for Random Assignment of Cases to Judges.

	Dependent Variable			
	Case Allowed		Judge Leniency	
	coeff.	s.e.	coeff.	s.e.
Age	0.0055***	(0.0009)	0.0003*	(0.0002)
Female	0.0140	(0.0095)	0.0004	(0.0018)
Married	0.0069	(0.0075)	0.0015	(0.0019)
Foreign born	-0.0277***	(0.0116)	0.0011	(0.0024)
High school degree	0.0126*	(0.0073)	0.0003	(0.0014)
Some college	0.0259	(0.0170)	-0.0005	(0.0033)
College graduate	-0.0974***	(0.0177)	0.0041	(0.0094)
One child	-0.0066	(0.0088)	-0.0011	(0.0019)
Two children	-0.0134	(0.0135)	0.0010	(0.0015)
Three or more children	-0.0343***	(0.0140)	0.0026	(0.0022)
Average indexed earnings	0.0000***	(0.0000)	0.0000	(0.0000)
Experience	0.0073***	(0.0008)	0.0000	(0.0001)
Mental disorders	0.0282*	(0.0147)	0.0013	(0.0059)
Musculoskeletal disorders	0.0651***	(0.0163)	0.0007	(0.0059)
F-statistic for joint significance	9.90		.75	
[p-value]	[.001]		[.720]	
N	14,893		14,893	

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Baseline estimation sample, consisting of parents who appeal an initially denied DI claim during the period 1989-2005 (see Section 3 for further details).. There are 79 different judges. Columns 1 and 3 display OLS estimates from separate regressions of whether a case is allowed or judge leniency, respectively, on appellant characteristics. F-statistics are obtained from OLS estimation on the combined set of applicant characteristics. All regressions include fully interacted year and department dummies. Characteristics of appellants are measured prior to the appeal. Number of children is the number under age 18, average indexed earnings is mean earnings for the last ten years prior to appeal and experience is number of years with positive earnings over this ten year period.

a judge's average allowance rate against this measure of judicial experience. While experienced judges appear to be slightly less lenient, experience accounts for only a small fraction of the total variation in allowance rates across judges. Other unobserved factors must be driving the underlying variation. It is important to recognize that as long as judges are randomly assigned, it does not matter why some judges are more lenient than others.

Using Judge Leniency as an Instrument

We use variation in DI receipt generated from the random assignment of appeal judges to estimate the intergenerational link in DI receipt. If we could randomly assign parents to a treatment group which gets DI and a control group which does not, then there would be no omitted variable bias, since the parent's (and child's) pre-assignment earning capacity, health and all other characteristics would, on average, be the same in the two groups. While we cannot implement this experiment, we can take advantage of the naturally occurring variation in the probability a parent will receive DI based on the judge which handles their appeal case. As we document below, some judges are systematically more lenient than others. Letting z_i be a judge's propensity to issue a lenient ruling, and assuming a linear model, the probability child i's parent will receive DI is:[14]

$$P_i^p = \alpha^p + \gamma^p z_i^p + \delta^p x_i^p + \varepsilon_i^p \tag{4}$$

Although we do not observe a judge's leniency directly, we can consistently estimate it by taking the average allowance rate in all other cases he or she has handled, as we did for Table 2.[15] Since judges are randomly assigned, their leniency will be uncorrelated with the error term in equation (1). This means we can use it as an instrumental variable in a standard two-stage least squares regression, where equation (4) is the first stage and a linear probability model of equation (1) is the second stage. The intuition behind this approach is that we only use the variation in parental DI which is driven by idiosyncratic differences in judge leniency to estimate the effect of parental DI receipt on their child. We can also estimate the reduced form effect by directly regressing P_i^c on z_i^p.

3 Data and Background

3.1 Data and Sample Restrictions

Our analysis employs several data sources that we can link through unique identifiers for each individual. Information on DI benefits comes from social security registers that contain complete records for all individuals who entered the DI program during the period 1967-2010. The data set includes information on the individual's work history and medical diagnosis,[16] the month when DI was awarded (or denied), and the level of DI benefits received. We link this information with administrative data from the hearing office on all appeals from 1989 to 2011. The data set contains information on dates of appeal and decision, the outcome of the appeal, and unique identifiers for both judges and applicants. We merge these data sets with administrative registers provided by Statistics Norway, using a rich longitudinal database that covers every resident from 1967 to 2010. For each year, it contains individual demographic information (including sex, age, and number of children), socio-economic data (such as years of education and earnings), and geograph-

[14] In this specification, we have omitted the grandparent's participation, P_i^g, which means that its effect will be a part of the error term.

[15] Although the instrument is pre-estimated, there is no need to adjust the standard errors of the IV estimates; such adjustments are necessary with generated regressors but not with generated instruments.

[16] Medical diagnoses are only available from year 2000 and onwards.

ical identifiers. The data contains unique identifiers that allow us to match spouses and parents to their children. The coverage and reliability of Norwegian registry data are rated as exceptional in international quality assessments (see Atkinson, Rainwater, and Smeeding 1995).

Our empirical analysis considers children of parents who appeal an initially denied DI claim.[17] Following Maestas, Mullen, and Strand (2013) and French and Song (2013), our baseline estimation excludes observations for which the assigned appeal judge has handled few cases (less than ten during the period 1989 to 2011). The reason for this sample restriction is to reduce the noise in our instrument. We further refine the sample to be appropriate for studying intergenerational transmission of DI receipt. We begin by restricting the sample to children whose parent's appeal decision was made made during the period 1989 to 2005. This sample restriction allows us to observe the behavior of the child for at least five years after appeal decision of the parent. We further exclude children whose parent were older than 55 years at the time he or she appealed. The reason for this age restriction is to avoid program substitution between DI and early retirement schemes.

In our main analysis, we restrict the sample to children who are age-eligible for DI (at least 18 years old) at the time of the parent's appeal decision. This age restriction allows us to observe participation behavior over time for a sizeable sample of children; the baseline sample consists of 14,893 parent-child observations and 79 different judges. One implication of the age restriction is that the baseline sample will be comprised of older children as compared to the unrestricted sample of appellants. Appendix Figure A.2 displays the age distribution of parents who appeal and the age distribution of their children. Because few parents with young children apply for DI, the baseline sample includes the typical parent-child links. In Section 6, we will nevertheless explore the impact of parental DI participation on an alternative, smaller sample of children who are under 18 at the time of the parent's appeal decision.

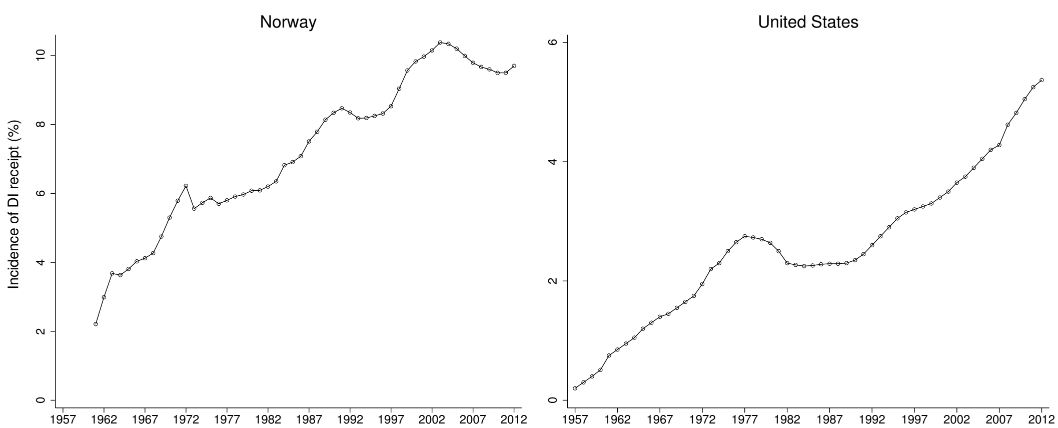
In Appendix Table A.1, we document the key characteristics of the sample of parents who apply for DI and our baseline sample of parents who appeal an initially denied DI claim. The parents who appeal are on average more likely to be female, less educated and foreign born, and have lower prior earnings and less work experience compared to the group of initial applicants. The children of parents who appeal tend to be less educated, but actually have slightly higher prior earnings compared to children of parents who initially apply for DI.

[17] Some parents have several denied DI claims over the period we consider. In such cases, we restrict our sample to the parent's first denied DI claim.

3.2 Institutional Background

There are a number of similarities and a few key differences between the DI systems in the U.S. and in Norway.[18] In both countries, DI is one of the largest transfer programs. However, the incidence of receipt of DI benefits is lower in the U.S. than in Norway. Figure 2 shows this distinction by displaying the evolution of DI in the two countries. Whereas the rate of DI receipt in a given year is consistently higher in Norway than in the U.S.,[19] the time trends are quite similar. From 1961 to 2012, the rate of receipt increased from 2.2 to 9.7 percent in Norway and from 0.8 to 5.4 percent in the U.S. While Norway's rate has leveled off at about 10 percent in recent years, the U.S. DI rate continues to rise and is projected to exceed 7 percent by 2018 (Burkhauser and Daly, 2012).

Figure 2: Trends in DI Receipt in Norway and the U.S.



Notes: U.S. trends based on Autor and Duggan (2006) for 1957-2005 and SSA Office of the Chief Actuary for 2006-2012. Norwegian trends based on SSA Statistical Supplements. Incidence of DI receipt defined as the percent of the relevant adult population receiving DI benefits (age 18-67 in Norway; age 25-64 in the US).

In both countries, the expansion of the DI rolls in recent decades appears to be driven by the liberalization of the screening process, which led to a rapid increase in the share of DI recipients suffering from difficult-to-verify disorders such as mental illness and musculoskeletal disease.[20] Because these are early-onset disorders with low age-specific mortality, DI recipients with such diagnoses tend to participate in the program for

[18]Our discussion of the U.S. system draws primarily on Autor and Duggan (2006), and pertains only to the SSDI program. More than 80 percent of non-elderly U.S. adults are insured against the risk of disabling physical or mental illness by SSDI. Our discussion of the Norwegian system is based on Kostøl and Mogstad (2013).

[19]The cross-country difference in DI coverage is unlikely to explain the entire discrepancy in the incidence of DI: although virtually all non-elderly adults are covered in Norway, more than 80 percent of all non-elderly adults are covered in the U.S. The remaining difference could be a function of underlying differences in screening stringency, the generosity of the programs, the frequency with which people apply for disability benefits or the health of the population. Milligan and Wise (2011) argue that differences in health are unlikely to explain much of the observed differences in DI rates across developed countries.

[20]See Autor and Duggan (2006) for a discussion of this phenomenon. In the U.S., the 1984 congressional reforms shifted the focus of screening from medical to functional criteria. In Norway, the medical eligibility criteria were relaxed earlier and more gradually.

relatively long periods. As a result, the DI exit rates have decreased in the last few decades. Appendix Figure A.3 displays the evolution of DI exit rates in the U.S. and Norway. In 1985, the yearly DI exit rate was approximately 12.1 percent in the U.S. and 10.4 percent in Norway. In both countries, this rate has trended steadily downward since that time and reached approximately 7 percent in 2004. As shown in Appendix Figure A.4, this decline has been driven both by a decrease in the fraction of DI recipients who reach retirement age and by a decrease in the fraction of DI recipients who die in any given year.

Another difference is that DI recipients in Norway tend to be somewhat older and to have slightly higher earnings prior to a disability award. Appendix Table A.2 report key characteristics of DI recipients in the U.S. and in Norway. One explanation for these differences in characteristics is that the U.S. SSDI program is less generous.[21] The differences in characteristics are, however, less pronounced than one might expect. For instance, almost 60 percent of DI recipients suffer from difficult-to-verify disorders (including mental illness and musculoskeletal disorders) in both the U.S. and Norway.

A third difference is that the appeal process plays a more important role in the U.S. than in Norway. In both countries, the disability determination process is a multi-step process, where cases that are difficult to judge are often denied at the initial application step. If the DI claim is initially denied, the individual may appeal the decision to the Court of Appeals where the appeals are reviewed by Administrative Law Judges. While 48 percent of the initially rejected applicants appeal in the U.S. (French and Song, 2013), only 25 percent of the initially rejected appeal in Norway. Appendix Table A.3 compares the characteristics of individuals who apply for DI and those who appeal an initially denied DI claim in the two countries. Both in the U.S. and in Norway, appellants are more likely to be younger, less connected to the labor market, and more likely to suffer from difficult-to-verify disorders, as compared to the the initial group of applicants. This suggests that in both countries the marginal applicants are often initially denied, and they are relatively likely to appeal.

4 Evidence on Intergenerational Welfare Transmission

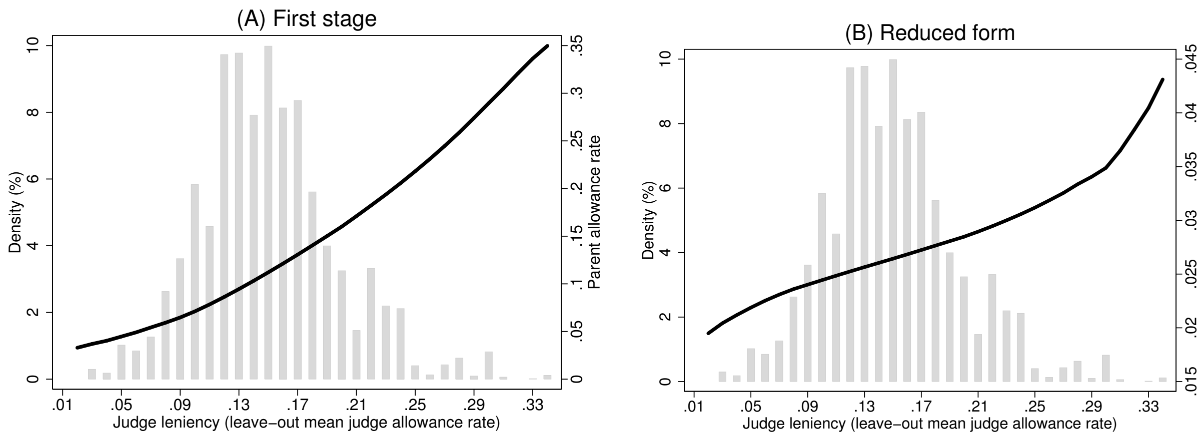
4.1 Graphical Evidence

We begin our presentation of results by providing a graphical representation of the IV approach in Figure 3. In the background of each graph is a histogram for the density of judge leniency, which captures the average

[21] For a typical DI recipient in Norway, Kostøl and Mogstad (2013) calculate the replacement rate would be 31 percent according to U.S. program rules and 58 percent according to Norwegian program rules. Factoring in health insurance coverage increases the effective replacement rate to over 50 percent in the U.S. In Norway, all citizens are eligible for health insurance through the Social Insurance system.

judge allowance rate in the other cases a judge has handled. We note the judge leniency measure is calculated from all cases the judge has ever handled, not just the cases in our estimation sample. On average, each judge has handled a total of 380 cases. The mean of the leniency variable is .15 with a standard deviation of .06. The histogram reveals a wide spread in judge leniency, with approximately 22% of cases allowed by a judge at the 90th percentile compared to approximately 9% at the 10th percentile.

Figure 3: Effect of Judge Leniency on Parents (First Stage) and Children (Reduced Form).



Notes: Baseline sample, consisting of parents who appeal an initially denied DI claim during the period 1989-2005 (see Section 3 for further details). There are 14,893 individual observations and 79 different judges. Panel (A): Solid line is a local linear regression of parental DI allowance on judge leniency. Panel (B): Solid line is a local linear regression of child DI receipt on their parent's judge leniency measure. All regressions include fully interacted year and department dummies. The histogram of judge leniency is shown in the background of both figures (top and bottom 0.5% excluded from the graph).

Panel A shows the effect of judge leniency on a parent's allowance rate. The graph is a flexible analog to the first stage equation (4), where we plot a local linear regression of actual parental allowance against judge leniency. The parental allowance rate is monotonically increasing in our leniency measure, and is close to linear. A one percentage point increase in the judge's allowance rate in other cases is associated with an almost one percentage point increase in the probability the parent's case is allowed. Panel B plots the reduced form effect of a parent's judge leniency measure against their child's DI participation, again using a local linear regression. The child's DI rate is monotonically increasing in the leniency measure as well. Approximately two and a half percent of children whose parents had a relatively strict judge (leniency measure =.09, the 10th percentile) are predicted to participate in DI five years later. This can be contrasted with roughly three percent of children whose parents had a relatively lenient judge (leniency measure = .22, the 90th percentile).

4.2 Intergenerational Transmission Estimates

We now turn to a regression based analysis. Column 1 in Table 3 reports first stage estimates which regress a dummy variable for whether a parent is allowed DI at the appeal stage on our judge leniency measure. We include fully interacted year and department dummies in the first column, but otherwise include no other controls. The coefficient implies that when a judge's allowance rate in the other cases he has handled goes up by 1 percentage point, the probability a parent will be allowed DI by that judge increases by 0.91 percentage points. This effect is not statistically different from one.

Table 3: Estimates of Intergenerational Welfare Transmission.

	First stage		Reduced form		IV		N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: *Child on DI five years after parent's appeal decision*							
Parent allowed DI	0.913***	0.868***	0.054**	0.052**	0.059***	0.059***	14,893
	(0.113)	(0.115)	(0.020)	(0.020)	(0.021)	(0.022)	
Additional controls?	NO	YES	NO	YES	NO	YES	
Dependent mean	0.11	0.11	0.03	0.03	0.03	0.03	
Panel B: *Child ever on DI after parent's appeal decision*							
Parent allowed DI	0.913***	0.868***	0.109***	0.104***	0.119***	0.120***	14,893
	(0.113)	(0.115)	(0.030)	(0.027)	(0.032)	(0.031)	
Additional controls?	NO	YES	NO	YES	NO	YES	
Dependent mean	0.11	0.11	0.08	0.08	0.08	0.08	

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

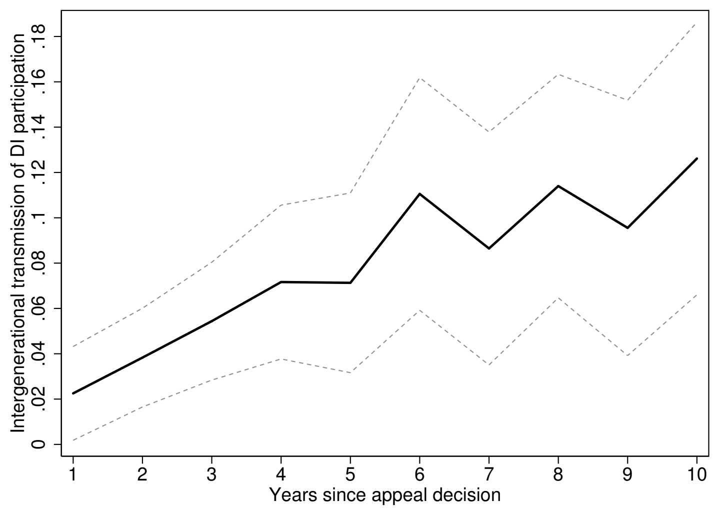
Notes: Baseline sample, consisting of parents who appeal an initially denied DI claim during the period 1989-2005 (see Section 3 for further details). There are 79 different judges. All regressions include fully interacted year and department dummies. Specifications with additional controls include a linear term for average indexed earnings and dummy variables for month of appeal, county of residence, age of parent and child, gender of parent and child, foreign born, marital status, number of children, education, labor market experience, and a number of medical diagnoses. The control variables are measured prior to the appeal. Number of children is the number under age 18, average indexed earnings is mean earnings for the last ten years prior to appeal and experience is number of years with positive earnings over this ten year period.

Panel A reports results for whether the child participates in DI within 5 years after the parent's appeal decision. Column 3 reports the reduced form estimate of a parent's judge leniency measure for this child outcome. The estimate of .054 implies that when judge leniency for a parent rises by 10 percentage points, a child's DI participation will rise by roughly one-half of a percentage point. This is a sizeable effect compared to the 3 percent average DI participation rate within five years for this sample. Column 5 takes the reduced form estimate of column 3 and divides it by the first stage estimate in column 1. Since the first stage is close to one, the reduced form and the IV estimates are very similar.

Panel B performs a similar exercise, but now looks at whether the child has ever been on DI after the parent's appeal decision. While every child is observed for at least five years after their parent's appeal decision, in this second panel some children will be observed for up to 21 years and, on average, the children are observed for 11 years. The unbalanced nature of this second panel affects the interpretation of the estimates, but it should not affect their validity given the nature of our instrument. The results suggest the long-run effects of a parent getting on to DI are larger than the short-run effects: the IV estimate rises to 12 percentage points in Panel B. While it is true that the mean of the dependent variable also increases in Panel B, the findings indicate that a parent's experience with the DI system is not just changing the timing of when their children participate in DI.

Figure 4 complements Table 3 by showing IV estimates for the intergenerational transmission of DI participation over time for a balanced panel. The estimates correspond to those in Table 3, except the graph restricts the sample to children observed for 10 years after their parent's appeal decision.[22] The effect grows substantially over time. Ten years after the court decision, the causal effect of a parent being allowed DI is a 12 percentage point increase in a child's DI take up.

Figure 4: Estimates of Intergenerational Transmission over Time.



Notes: Baseline sample restricted to parents who appeal an initially denied DI claim during the period 1989-2000, so as to have a balanced 10 year sample. There are 9,143 individual observations and 50 different judges. The figure displays separate IV estimates of intergenerational transmission 1 to 10 years after the parent's appeal decision. The specifications mirror column 6 of Table 3. Dashed lines represent 95 percent confidence intervals (clustered at the judge level).

Lastly, we shift attention to how a parent's DI participation affects the probability that their children subsequently apply for DI. Appendix Figure A.5 shows IV estimates for child DI application over time based on the balanced panel. These results mirror closely the estimates for DI participation. The effect on DI

[22]The first stage estimate for this sample is 0.988 with a standard error of 0.146.

application grows substantially over time. Ten years after the court decision, the causal effect of a parent being allowed DI at the appeal stage is a 14 percentage point increase in a child's DI application rate. Given the similarity in the estimates, we will focus on children's DI participation in the remainder of the paper.

4.3 Internal Validity

In order for judge leniency to be a valid instrument, appellants' assignment to judges must be uncorrelated with case characteristics (conditional on fully interacted year and department dummies). This amounts to an assumption of random assignment among the pool of judges within each department. Table 2 provided strong empirical support for the claim that the DI system in Norway randomly assigns appeal judges. The even numbered columns of Table 3 explore what happens if a large set of control variables are added to the regressions. If judges are randomly assigned, the addition of these control variables should not significantly change the estimates, as both parental and child characteristics should be uncorrelated with judge leniency. As expected, the coefficients do not change appreciably. As a final test of randomization, we examine whether the likelihood of children receiving sickness pay prior to the parents' appeal is correlated with judge leniency. Before going onto DI, individuals usually participate in the sickness program; correlation between our instrument and children's pre-determined participation rate in this program would therefore raise concerns about compliance with the random allocation procedure. It is reassuring to find that child participation in the sickness program is not statistically related to the leniency of the judge assigned to their parent's case.[23]

While random assignment of cases to judges is sufficient for a causal interpretation of the reduced form estimates, the IV estimates require two additional assumptions. The first is that the parent's draw of a judge affects the child's DI participation only through the judge's allowance or denial decision. This exclusion restriction implies that parental DI allowance is the unique channel for any causal effects of judge leniency. One attractive feature of the process in Norway makes this exclusion restriction likely to hold: the appeal is presented in writing, so there is never any personal contact between the judge and those who appeal. What parents and children observe is the allowance or denial decision of the judge. A possible caveat is that the appeal processing time could differ systematically by the leniency of the judge (see e.g. Autor, Maestas, Mullen, and Strand (2011)) and that this could affect a child's decision to apply for DI. To examine this, we calculated judge processing time based on the residual average processing time in the other cases a judge

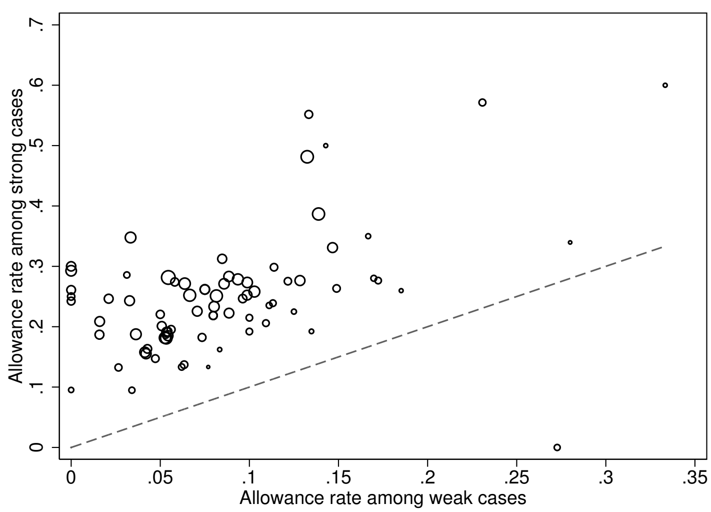
[23] The regression coefficient of judge leniency on child's participation in the sickness program is 0.009 (s.e. = 0.053). This point estimate is small compared to the sample mean: 23 percent of children had received sickness pay at some point prior to their parent's appeal.

has handled after controlling for a fully interacted set of time and department dummies in a regression. It is reassuring to find that our instrument, judge leniency, and judge processing time are virtually uncorrelated. Moreover, the first row of Table 4 shows that the IV estimates do not change appreciably if we control for judge processing time in the first and second stages.

The final assumption needed for a causal interpretation of the IV estimates is monotonicity of judge's appeal decisions. In our setting, the monotonicity assumption is that cases allowed by a strict judge would also have been allowed by a more lenient judge, and similarly that cases denied by a lenient judge would also have been denied by a stricter judge. One testable implication of the monotonicity assumption is that the first stage estimates should be non-negative for all subsamples. Appendix Table A.5 provides separate first stage estimates based on characteristics of the parent and the child. These estimates are consistently positive and sizeable, in line with the monotonicity assumption.

Another test of the monotonicity assumption can be performed by seeing whether individual judges make similar decisions about the relative strength of a case, a prediction we explore in Figure 5. To conduct this test, we first run a regression similar to column 1 in Table 2 to predict the probability a case will be allowed.[24] Based on these predictions, we create two equally-sized groups: appellants who have a strong case (an above median allowance probability) and those who have a weak case (a below median allowance probability). For

Figure 5: Individual Judge Allowance Rates in Strong vs. Weak Cases.



Notes: Each circle plots an individual judge's average allowance rate in their stronger versus weaker cases. There are 79 judges. A strong (weak) case is defined as having a predicted allowance probability above (below) the median (see text for details). Circle size is proportional to the t-statistic for the difference in means. Dashed line is the 45 degree line.

[24] For increased precision in this preliminary step, we use all observations from judges handling at least 10 cases, whether or not the cases appear in our baseline estimation sample. This results in a sample of 107 judges who handle 28,392 appeal cases.

each judge in our baseline estimation sample, we take their caseload and divide it up into these two groupings. If monotonicity holds, each individual judge should be allowing their strong cases at a higher rate than their weak cases. Indeed, this is the case for all but one judge in our sample, as graphed in Figure 5. That figure plots each judge's allowance rate in weak cases versus their allowance rate in strong cases, with almost all observations lying above the 45 degree line.[25]

Lastly, Table 4 reports the results from several specification checks, all of which support our main findings. In the second row, we limit the sample to the period when there was just one department, rather than four departments handling appeals. While the standard errors go up somewhat, the results are similar. The third and fourth row show the results are robust to adding in fully interacted year, month and department dummies or excluding parents who die. In our baseline analysis, we excluded judges who handle less than 10 cases. The fifth and sixth rows demonstrate that including these judges does not change the estimates appreciably, and neither does excluding judges who handle less than 50 cases. The final row considers an alternative handling of remanded cases. In our baseline analysis, we code a remanded case as rejected (see footnote 9). If we instead code remanded cases as allowed or denied based on its eventual outcome after it is reconsidered by the DDS case worker with updated information, the results are quite similar.

Table 4: Specification Checks for Intergenerational Welfare Transmission Estimates.

	Child on DI five years after parent's appeal decision		
	First stage	IV	N
With judge processing time	0.856***	0.059***	14,893
	(0.107)	(0.022)	
One Department (pre-1998)	1.090***	0.054**	5,589
	(0.185)	(0.025)	
Month-department controls	0.773***	0.065**	14,893
	(0.128)	(0.029)	
Exclude parents who die	0.859***	0.067***	14,504
	(0.113)	(0.023)	
Include judges $<$ 10 cases	0.855***	0.059***	14,897
	(0.116)	(0.023)	
Exclude judges $<$ 50 cases	0.944***	0.055**	14,758
	(0.111)	(0.021)	
Alternative coding of remand	0.802***	0.064***	14,893
	(0.103)	(0.022)	

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Baseline sample (see Table 3). The specifications mirror the IV model with additional controls reported in column 6 of Table 3.

[25] While the grouping into strong and weak cases is based on a regression using all cases (including a judge's own cases), since there are a reasonably large number of judges and cases, this should not overly influence the results. The one outlier below the 45 degree line in the graph represents a judge who handled only 12 cases.

4.4 Interpreting the IV estimates

Our IV estimates should be interpreted as a local average treatment effect (LATE) for children whose parents would have received a different allowance decision had their case been assigned to a different judge. Our instrument picks out these complier children, whose parents are on the margin of program entry. To better understand this LATE, we use the methods of Imbens and Rubin (1997) and Abadie (2003) to count the compliers, estimate their potential outcomes and explore their observable characteristics.

We first calculate the number of always takers, never takers and compliers in our sample. These compliance types are usually defined in the context of binary instruments, whereas judge leniency is a continuous instrument. We therefore look at the allowance rates for the "most lenient" and the "strictest" judges. Our first stage coefficient, combined with these allowance rates, is informative about the number of appellants who would have received a different allowance decision had their case been assigned to a different judge.[26] We estimate that these compliers make up 23% of our sample. Because of monotonicity, the share of parents that would be allowed DI regardless of the judge assigned to their case is given by the probability of allowance for the strictest judge. These always takers make up 13% of the sample. The remaining 64% of our sample are never takers who would not be allowed DI no matter which judge was assigned to their case. We also characterize compliers by observable characteristics in Appendix Table A.5. The most distinctive feature of these marginal applicants is their educational attainment: Sixty-five percent of the compliers have low education, while their fraction of the entire sample is only 56 percent.

We next estimate the potential participation rates behind the LATE. Our IV estimates tells us that the probability a complier child has ever been on DI after their parent's appeal increases by 12 percentage points if the parent is allowed DI. A natural question would be, how many complier children would have been on DI if their parents had been denied DI? We can recover this potential outcome by combining (i) estimates of the mean child outcomes of the always takers and the untreated with (ii) our estimates of the shares of always takers, never takers and compliers.[27] We estimate that only 1 percent of the complier children would have

[26] For ease of discussion, consider the case where the first stage, equation (4), has no covariates. The share of compliers is given by $\pi_c \equiv Pr(P^p = 1 \mid z_i^p = \overline{z}) - Pr(P^p = 1 \mid z_i^p = \underline{z}) = \gamma^P(\overline{z} - \underline{z})$, where $\overline{z}$ and $\underline{z}$ denote the maximum and minimum values of the instrument. The share of always takers is given by $\pi_a \equiv Pr(P^p = 1 \mid Z = \underline{z}) = \alpha^P + \gamma^P \underline{z}$ and the share of never takers is given by $\pi_n \equiv Pr(P^p = 0 \mid z_i^p = \overline{z}) = 1 - \alpha^P - \gamma^P \overline{z}$. To estimate these quantities we use estimates of the first stage coefficients (Table 3, column 2) and the average leniency measures for the top one percentile (most lenient) and bottom one percentile (strictest) of judge leniency.

[27] Let $P^c(0)$ and $P^c(1)$ denote the potential DI participation of a child based on whether the parent is denied or allowed DI, respectively. The LATE is given by the difference in potential participation of complier children with parents denied vs. allowed DI, $E(P^c(1)|c) - E(P^c(0)|c)$. Continuing with the notation developed in footnote 26, the potential participation of children of never takers denied DI is given by $E(P^c(0)|n) = E(P^c|P^p = 0, z_i^p = \overline{z})$, while the potential participation of children of always takers allowed DI is $E(P^c(1)|a) = E(P^c|P^p = 1, z_i^p = \underline{z})$. The potential participation of complier children can be inferred from $E(P^c \mid P^p = 0, z_i^p = \underline{z}) = \frac{\pi_c}{\pi_c+\pi_n} E(P^c(0) \mid c) + \frac{\pi_n}{\pi_c+\pi_n} E(P^c(0) \mid n)$ and $E(P^c \mid P^p = 1, z_i^p = \overline{z}) = \frac{\pi_c}{\pi_c+\pi_a} E(P^c(1) \mid c) + \frac{\pi_n}{\pi_c+\pi_a} E(P^c(1) \mid a)$.

ever been on DI after their parent's appeal if their parents had been denied DI. But if these same parents were awarded DI instead, the DI participation rate for their children would be 13 percent.

The fact that compliers are a fairly small and distinct group helps explain why our IV estimates exceed the OLS estimates reported in Table 1. It also means that we need to be cautious in extrapolating our results to the population at large. However, the intergenerational link among compliers is particularly relevant for policy, since reforms aimed at stemming the rise in DI will likely have the largest effect on applicants on the margin of program entry. Furthermore, in both Norway and the U.S., the rise in DI rolls in recent decades appears to be primarily driven by a more liberal screening of marginal applicants who are often initially denied and relatively likely to appeal (Autor and Duggan, 2006, Kostøl and Mogstad, 2013).

Lastly, we explore the counterfactual labor and educational outcomes for children, i.e., what would have happened if a parent had not been allowed on DI? Consistent with the impact on children's use of DI, we find that parental DI participation decreases the probability that a child will be employed or pursue higher education. Examining child outcomes five years after their parent's appeal, Table 5 shows that a parent's DI receipt causes employment to drop by 13 percentage points. While we do not estimate the drop in full-time work or college completion with the same precision, both estimates suggest a sizeable drop in these child outcomes as well.

Table 5: Effect of Parent's DI Participation on Child Labor and Educational Outcomes.

Dependent variable	Reduced form	IV	Dep. mean	N
Employment	-0.117**	-0.134**	0.57	14,893
	(0.054)	(0.063)		
Full-time work	-0.066	-0.077	0.42	14,893
	(0.078)	(0.088)		
College degree	-0.070	-0.080	0.25	14,893
	(0.058)	(0.067)		

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Baseline sample (see Table 3). Specifications mirror column 6 of Table 3. Employment is defined as working more than 4 hours a week, full-time work as more than 30 hours a week, and college degree as having completed college by 2010. Labor outcomes are measured five years after parent's appeal decision.

5 Policy Simulation

Our results provide strong evidence that welfare use in one generation causes welfare use in the next generation. These intergenerational effects could have important implications for the evaluation of welfare reforms, as any changes will affect not only the current generation, but also have spillover effects on their children.

In this section, we simulate the total reduction in DI participation from a policy which makes the screening process more stringent. This simulation makes clear that accounting for intergenerational effects is key to make accurate projections of post-reform participation rates and program costs.

We consider a policy change which makes all judges one-fifth of a standard deviation less likely to allow an appeal (a change of 0.012 in our judge leniency variable), a policy which conceivably could be achieved by instructing judges to be stricter in their rulings. This change translates into the average judge being approximately 10 percent less likely to grant an appeal. Our simulation focuses on how this policy change would affect the participation rates of parents and children in the balanced 10 year sample. In particular, we abstract from any behavioral responses among parents at the margin of applying for DI; we further assume no changes in the screening by case examiners at the initial application step.

There are two components to the total reduction in DI from the policy change: the direct effect on parents, and the indirect effect on children. To calculate the direct effect on parents, we regress parental DI participation in a given year on judge leniency, and multiply this estimated coefficient by (minus) one-fifth of a standard deviation. We perform a similar calculation for children, regressing child DI participation in a given year on their parent's judge leniency measure, and multiply this estimated coefficient by (minus) one-fifth of a standard deviation. We then calculate how much these direct and indirect effects would lower DI participation over time. Table 6 displays the estimated coefficients of judge leniency on child and parental DI participation in every second year. The effect of judge leniency on child DI participation grows substantially

Table 6: Effect of Judge Leniency on Child and Parent DI Participation Over Time.

	Years since court decision				
	2 years	4 years	6 years	8 years	10 years
	Panel A: DI Child				
Estimate	0.038***	0.071***	0.109***	0.113***	0.125***
Standard error	(0.011)	(0.017)	(0.025)	(0.027)	(0.036)
Dependent mean	0.010	0.018	0.030	0.046	0.063
	Panel B: DI Parent				
Estimate	0.688***	0.630***	0.431***	0.379***	0.259***
Standard error	(0.134)	(0.127)	(0.112)	(0.102)	(0.090)
Dependent mean	0.385	0.508	0.581	0.631	0.669

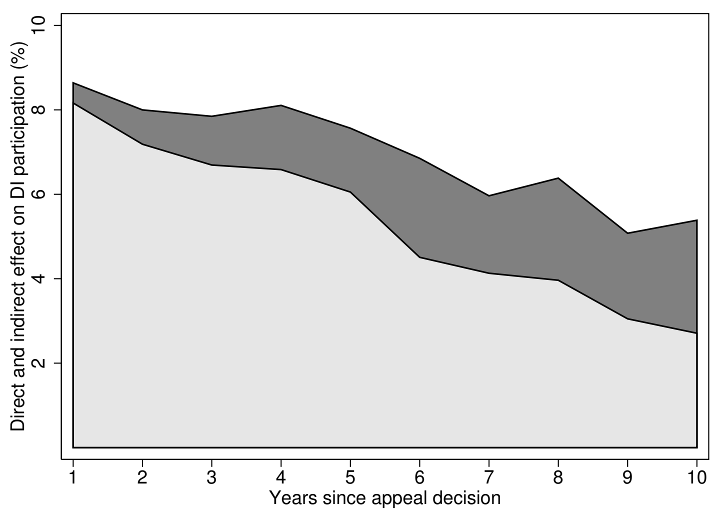
***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Balanced 10 year sample created by restricting the baseline sample to parents who appeal an initially denied DI claim during the period 1989-2000. There are 9,143 individual observations and 50 different judges. Panel A regresses child DI participation in a given year on their parent's judge leniency. Panel B regresses parent DI participation in a given year on parent's judge leniency. Specifications mirror column 6 of Table 3.

over time. In contrast, the effect of judge leniency on parental DI participation shrinks over time. This is in part because some initially rejected parents re-apply and are awarded DI and in part because some parents reach early retirement age and exit DI.

Using the estimates in Table 6, we graph the results of the policy simulation in Figure 6. In the first year after the court decision, making judges one-fifth of a standard deviation less likely to allow an appeal reduces DI participation by almost 9 percent. Most of this initial reduction can be attributed to the direct effect on parents, as there is little opportunity for children to learn and respond to their parent's DI experience. Over time, however, the direct effect of tightening the appeals process shrinks; by year 10, the direct effect of the policy change results in a 3 percent drop in DI rolls. In contrast, the indirect intergenerational effect grows over time. After ten years, the increase in children's participation accounts for a 3 percent reduction in the DI rolls. Taken together, these results show that in the first years after making the DI program more stringent, almost all of the drop in participation is due to the fact that fewer parents are being allowed DI; but 10 years later, almost half of the reduction in DI is accounted for by the reduced participation of the children of the original applicants.

Figure 6: The Effect of Tightening the Screening Process on Parents and Their Children.



Notes: Balanced 10 year sample (see Table 6). The figure display the direct and indirect effects of tightening the screening process by one-fifth of a standard deviation. Light gray area is the direct effect on parents' participation. Dark gray area is the indirect effect on their children due to the intergenerational transmission of DI use. The estimated coefficients underlying this graph are shown for every second year in Table 6.

This simulation makes clear that failing to account for intergenerational effects will provide misleading projections of post-reform participation rates and program costs. To translate the participation patterns shown in Figure 6 into cost terms, we calculated the net present value of the simulated policy change for parents and children over time, based on average DI benefit amounts and assuming a 3 percent annual discount rate. Making judges approximately 10 percent stricter decreases the net present value of program expenditures after 10 years by roughly 8 percent. Two thirds of this cost reduction is due to fewer parents being on DI. But one third of the reduction is due to the fact that fewer children participate in DI as well. If one were to extrapolate past ten years by assuming that there are no further changes in DI take up among parents or children after year 10, and that parents and children stay on DI until they reach retirement at age 67, the contribution of children to total costs is even more important. Almost 60 percent of the reduction in total costs is now accounted for by the reduction in children's participation. This is due to the fact that children entering DI have many years left before retirement, while parents are older and age out of the system sooner.

6 Mechanisms

6.1 Costs and Benefits of DI Allowance for Parents

In this section, we estimate the effect of being allowed DI at the appeal stage on parent's subsequent labor market outcomes, benefit payments and income. This information is useful in understanding not only the behavioral responses of parents to the appeal decision, but also as background for understanding the type of information likely to be transmitted from parent to child about the costs and benefits of DI allowance.

Table 7 estimates how DI receipt affects parent's labor outcomes, using our judge leniency measure as an instrument. We find that being allowed DI at the appeal stage is associated with a sizeable drop in the fraction of parents receiving more earnings than the substantial gainful activity threshold.[28] Mean labor income and the probability of full-time work also fall substantially for parents allowed DI, as expected due to the program's eligibility rules. As reported in the table, we estimate that 5 years after the appeal, DI receipt causes parental earnings to decrease by almost \$8,000 and the probability of labor income being above the SGA threshold (full-time work) to decrease by 18.9 (12.1) percentage points.

Two studies using U.S. data and a similar research design have looked at how DI receipt affects labor supply. Exploiting the leniency of appeal judges in the U.S., French and Song (2013) find that DI receipt reduces earnings by \$4,915 and the probability of earning more than the SGA threshold by 14 percentage points five years later. Maestas, Mullen, and Strand (2013) exploit the leniency of initial examiners in the

[28] Earnings are not allowed to exceed this threshold once on DI. In 2005, the threshold was approximately \$10,000 annually.

Table 7: Effect of DI Allowance on Parent Labor Outcomes.

Dependent variable	Reduced form	IV	Dep. mean	N
Labor Income > SGA	-0.164**	-0.189**	0.31	14,893
	(0.072)	(0.079)		
Labor Income ($)	-6,924***	-7,980***	10,071	14,893
	(2507)	(2879)		
Full-time work	-0.105**	-0.121***	0.05	14,893
	(0.043)	(0.047)		

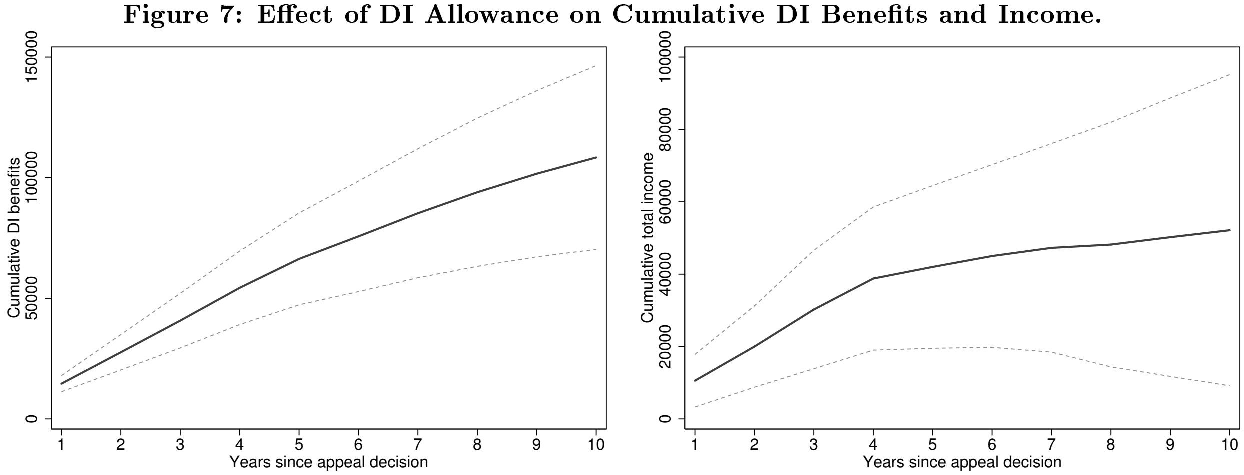
***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Baseline sample (see Table 3). Specifications mirror those of column 6 of Table 3. The substantial gainful activity level (SGA) equalled approximately $10,000 in 2005. Labor income converted to U.S. dollars (using the exchange rate NOK/$ = 6) and adjusted for inflation to 2005 levels. Full-time work is defined as working more than 30 hours a week.

U.S. and find that DI receipt reduces earnings by $3,800-4,600 and the probability of exceeding the earnings threshold by 18-19 percentage points two years later. By way of comparison, our labor supply effects for Norway are quite similar to those found for the U.S, which indicates that the link between DI receipt and labor market participation is comparable across the two countries.

Figure 7 complements Table 7 by showing IV estimates for the effect of DI allowance on DI benefits and total income (equal to labor income plus DI benefits) for the balanced sample. The estimates correspond to those in Figure 4, except that we now look at cumulative income and benefits up to 10 years after parent's appeal decision. We find an increasing and concave time pattern in the effect of being allowed DI at the appeal stage on cumulative benefits and total income. In the first years after the court decision, there is a

Figure 7: Effect of DI Allowance on Cumulative DI Benefits and Income.



Notes: Balanced 10 year sample (see Table 6). Specifications mirror those in Table 3. Solid lines are the yearly estimated effect of a parent's DI allowance, instrumented with judge leniency on DI benefits (first graph) and total income (DI benefits + labor income, second graph) accumulated since the appeal decision. Dashed lines represent 95 percent confidence intervals (clustered at the judge level). All monetary figures are reported in U.S. dollars (using the exchange rate NOK/$ = 6) and adjusted for inflation to 2005 levels.

large gain in DI benefits for the allowed parents. However, many parents who are initially denied re-apply and are eventually allowed DI (see Table 6). Indeed ten years after the appeal decision around two-thirds of these initially denied parents are receiving DI benefits. There are, however, large income losses to the initially denied parents for a number of years, especially during re-application, because if an applicant demonstrates that they have substantial earnings capacity, they are less likely to be allowed on to DI.

6.2 Intergenerational Transmission Channels

As emphasized by Moffitt (1992), there are at least three pathways which could drive a welfare culture within families: (i) parents on welfare could lessen the stigma of participation, (ii) parents on welfare could provide more information about welfare programs and less information about employment, and (iii) parents on welfare could invest differentially in child development. Each of these pathways indicates that it is the parent's experience with welfare programs that creates an intergenerational link. Although there are few previous studies to guide us on the role these channels play, the detailed nature of our data allows us to provide some suggestive pieces of evidence.

Parental Investments

Since we look at children who are at least 18 years old at the time of their parent's appeal decision, our estimates cannot be driven by differential parental investments in childhood or adolescence. Yet it could be that DI participation makes parents invest differentially in adult children. In Table 8, however, we show the causal intergenerational relationship remains strong even if we exclude children who live at home with their parents or focus on children who are at least 25 years old and tend to have completed their schooling.

Table 8: Intergenerational Welfare Transmission by Age and Living Arrangement of Child.

	Five years after parent's appeal			
	Reduced form	IV	Dep. mean	N
Child living away from home	0.076**	0.080***	0.03	8,652
	(0.031)	(0.031)		
Child at least 25 years of age	0.075**	0.073**	0.03	6,562
	(0.029)	(0.029)		

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Specifications mirror column 6 of Table 3. Child residency is determined based on whether a child has a different address from their parent one year prior to the parent's appeal.

While these results suggest that parental investment is not explaining our estimates of intergenerational welfare transmission, this channel could still be important for young children who live at home and grow up with a parent on DI. To examine this, Table 9 looks at an alternative, smaller sample of children who are under 18 (on average 13) at the time of their parent's appeal decision. We find that parental DI participation substantially increases the probability that their children will subsequently apply for DI. While we do not estimate the impact on child DI participation as precisely (p-value of 0.12), the estimates suggest a sizeable increase in this outcome as well. Taken together, the findings in Tables 8 and 9 suggest that differential investments by parents is not a key reason for the existence of family welfare cultures, at least in the context of DI.

Table 9: Effect of Parent DI Allowance Granted when a Child is Young (Less than Age 18).

	First stage	Reduced form	IV	N
Panel A: *Child applied to DI ten years after parent's appeal decision*				
Parent allowed DI	0.878***	0.084*	0.096**	4,172
	(0.138)	(0.043)	(0.045)	
Dependent mean	0.08	0.03	0.03	
Panel B: *Child on DI ten years after parent's appeal decision*				
Parent allowed DI	0.878***	0.062	0.071	4,172
	(0.138)	(0.042)	(0.046)	
Dependent mean	0.08	0.03	0.03	

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Sample is restricted to children of parents appealing their first denied case during the period 1989-2000. Furthermore, the children are restricted to be at most 17 years of age the year of their parent's appeal decision. The age range of children at the time of their parent's decision is 8-17, and the mean age ten years after is 23. Specifications mirror column 6 of Table 3.

Stigma versus Learning

The evidence against parental investments leaves us with two main competing explanations for our findings: a reduction in stigma or learning from increased information. Stigma can be thought of as a consumption complementarity in DI; if a parent is on DI, a child's utility of taking up DI could be higher. Information transmission happens because a child observes their parent's entire experience with the DI system and updates their expectations about the relative costs and benefits of applying for DI. Without data on subjective expectations and individual information sets, it is difficult to distinguish between stigma and learning. However, we find three pieces of suggestive evidence against the hypothesis that our findings are driven by a drop in social stigma resulting from parental DI use.

The first piece of evidence relies on the fact that all parents in our sample make it to the appeal stage. To qualify for DI benefits, a worker must demonstrate they have a physical or mental health impairment which substantially reduces their ability for gainful employment. Before applying for DI, a worker usually participates in the long-term sickness program, and if deemed able, in an employment rehabilitation program. Combined, these requirements mean the parents in our sample have already signaled a limited attachment to the labor market, likely been on other social assistance programs, applied for DI, and appealed an initial DI rejection. As a consequence, there is limited scope for a reduction in stigma arising from parents being allowed instead of denied DI at the appeal. For the same reason, our estimates are not likely to be driven by information transmission about how to initially apply to the DI program.

The second piece of evidence considers DI in the broader context of all welfare programs. Learning and information transmission from a parent's DI experience should be largely specific to the DI program, with little information which can be transferred to other welfare programs. In contrast, social stigma could relate more broadly to the take up of any welfare program, as argued by Blundell and Macurdy (1999). In our setting, if a parent participates in DI and stigma matters, it is likely the stigma of participating in other welfare programs would also go down somewhat.

Table 10 reports estimates of the effect of parental DI allowance on child participation in various welfare programs. As before, we use judge leniency as an instrument for parental DI allowance. For comparison, the first specification copies our baseline estimates for a child's DI participation, which are large and statistically significant. The second specification regresses a child's participation in Norway's social assistance program on their parent's DI allowance. This program is considered a last-resort safety net and there are no clear rules regarding eligibility or benefit amounts, with discretion being left to the local social worker. Appendix Figure A.6 reports survey evidence showing that participation in this program is highly stigmatized. Yet, both the reduced form and the IV estimates are small and statistically insignificant. The close to zero estimates are unlikely to reflect benefit substitution, as the correlation between DI and social assistance are slightly positive both in our sample (correlation = 0.07) and in the population at large (correlation = 0.10). When we look at whether a parent's DI experience affects the likelihood a child receives any other type of cash transfer (a broad measure which captures UI payments, housing benefits, etc.), we similarly find no effect. These findings do not support a model of general social stigma since such a model would have predicted increases in other forms of child welfare participation after a parent is allowed DI. Instead, we only find effects for a child's DI participation, which fits with a model of program-specific information being transferred from parent to child.

Our final piece of evidence relies on how DI take up rates for both parents and their children vary over time. The time patterns in Table 6 have different predictions for the stigma and learning channels. If

Table 10: Effect of Parent DI Allowance on Child Participation in Various Welfare Programs.

	Five years after parent's decision			
	Reduced form	IV	Dep. mean	N
Child on disability insurance	0.052**	0.059***	0.03	14,893
	(0.020)	(0.022)		
Child on social assistance (i.e., traditional welfare)	-0.008	-0.009	0.10	14,893
	(0.046)	(0.053)		
Child receiving other cash transfer	-0.009	-0.011	0.61	14,893
	(0.062)	(0.071)		

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Baseline sample (see Table 3). Specifications mirror those of column 6 of Table 3. Social assistance is a means-tested program for individuals with very low income. Other cash transfers includes unemployment benefits, benefits for single parents, child allowances, housing benefits, etc.

social stigma is the primary channel, this means that a child's utility of DI depends directly on a parent's participation. As more initially denied parents transition on to DI, the gap in stigma between the treatment group (the initially allowed parents) and the control group (the initially denied parents) should shrink. This predicts that the effects on children's participation should fade out over time as more denied parents get on DI. But in actuality, the estimated effect of parent's DI participation on a child's DI participation increases over time.[29] By comparison, a model in which children learn from their parent's cumulative experience fits the time patterns of DI participation for both parents and children. In such a model, children first learn about the initial rejection of a parent's DI application, and then as initially denied parents begin to re-apply for DI, their children additionally learn the process is time consuming, risky and increasingly costly (see Figure 7).[30]

7 Conclusion

As Black and Devereux (2011) conclude in their recent Handbook of Labor Economics chapter, despite large intergenerational correlations in welfare use, there is little evidence on the causal relationship. This paper provides strong evidence that such a causal link exists in the context of the DI program. The key to our identification approach is that judges are randomly assigned to DI applicants whose cases are initially

[29] While it is true children are getting older as time goes on, Appendix Table A.6 shows that the increasing effect on children's DI participation over time holds up even if we re-weight the sample so as to keep the distribution of child ages fixed.

[30] Of course, it could also be argued that a child with rational expectations could predict the costs and benefits of re-application, in which case there should be no change in a child's DI participation over time. It is also possible that seeing a denied parent eventually get on to DI encourages a child to apply for benefits, so that the gap in child DI participation should shrink over time. This would be a case of positive information being revealed over time. Neither of these information stories is consistent with the time pattern of effects in Table 6.

denied and some appeal judges are systematically more lenient. Using this random variation, we find strong evidence of a welfare culture, where welfare use in one generation causes welfare use in the next generation. When a parent is allowed DI because of a lenient judge, their child's participation over the next five years increases by 6 percentage points, an effect which grows over time. The detailed nature of our data allows us to explore the mechanisms behind the causal intergenerational relationship; we find suggestive evidence against stigma and parental investments and in favor of children learning from a parent's experience.

Our findings have important implications for the evaluation of welfare reforms, as any changes will affect not only the current generation, but also have spillover effects on their children. We use our estimates to simulate the total reduction in DI participation from a policy which makes the screening process more stringent. This simulation makes clear that accounting for intergenerational effects is key to make accurate projections of post-reform participation rates and program costs.

References

Abadie, A. (2003): "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, 113(2), 231–263.

Antel, J. J. (1992): "The Intergenerational Transfer of Welfare Dependency: Some Statistical Evidence," *The Review of Economics and Statistics*, 74(3), 467–473.

Atkinson, A., L. Rainwater, and T. Smeeding (1995): "Income Distributions in OECD countries: Evidence from the Luxembourg Income Study," *OECD Publications and Information Center*.

Autor, D. (2011): "The Unsustainable Rise of the Disability Rolls in the United States: Causes, Consequences, and Policy Options," *NBER Working Paper*, No. 17697.

Autor, D., N. Maestas, K. Mullen, and A. Strand (2011): "Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants," *University of Michigan Working Paper*, No. 258.

Autor, D. H., and M. G. Duggan (2003): "The Rise in the Disability Rolls and the Decline in Unemployment," *Quarterly Journal of Economics*, 118(1), 157–205.

——— (2006): "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding," *Journal of Economic Perspectives*, 20(3), 71–96.

Beaulieu, N., J.-Y. Duclos, B. Fortin, and M. Rouleau (2005): "Intergenerational Reliance on Social Assistance: Evidence from Canada," *Journal of Population Economics*, 18, 539–562.

Bertrand, M., E. F. P. Luttmer, and S. Mullainathan (2000): "Network Effects and Welfare Cultures," *The Quarterly Journal of Economics*, 115(3), 1019–1055.

Bird, A. (2007): "Perceptions of Epigenetics," *Nature*, 447(24), 396–398.

Bjorklund, A., M. Lindahl, and E. Plug (2006): "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *The Quarterly Journal of Economics*, 121(3), 999–1028.

Black, D., K. Daniel, and S. Sanders (2002): "The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust," *The American Economic Review*, 92(1), 27–50.

Black, S., and P. Devereux (2011): "Recent Developments in Intergenerational Mobility," chap. 16, pp. 1487–1541. Handbook of Labor Economics, Elsevier.

Blundell, R., and T. Macurdy (1999): "Labor Supply: A Review of Alternative Approaches," chap. 27, pp. 1559–1695. Handbook of Labor Economics, Elsevier.

Borghans, L., A. Gielen, and E. Luttmer (2012): "Social Support Substitution and the Earnings

Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform," *NBER Working Paper*, No. 18261.

BOUND, J. (1989): "The Health and Earnings of Rejected Disability Insurance Applicants," *American Economic Review*, 79(3), 482–503.

BRATBERG, E. (1999): "Disability Retirement in a Welfare State," *Scandinavian Journal of Economics*, 101(1), 97–114.

BRATBERG, E., Ø. A. NILSEN, AND K. VAAGE (2012): "Is Recipiency of Disability Pension Hereditary?," IZA Discussion Papers 6508, Institute for the Study of Labor (IZA).

BURKHAUSER, R. V., AND M. C. DALY (2012): "Social Security Disability Insurance: Time for Fundamental Change," *Journal of Policy Analysis and Management*, 31(2), 454–461.

CAMPOLIETI, M. (2004): "Disability Insurance Benefits and Labor Supply: Some Additional Evidence," *Journal of Labor Economics*, 22(4), 863–889.

CAMPOLIETI, M., AND C. RIDDELL (2012): "Disability Policy and the Labor Market: Evidence from a Natural Experiment in Canada, 1998-2006," *Journal of Public Economics*, 96, 306–316.

DOYLE, J. (2007): "Child Protection and Child Outcomes: Measuring the Effects of Foster Care," *The American Economic Review*, 97(5), 1583–1610.

——— (2008): "Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care," *Journal of Political Economy*, 116(4), 746–770.

DUNCAN, G., M. HILL, AND S. HOFFMAN (1988): "Welfare Dependence Within and Across Generations," *Science*, 239(4839), 467–471.

FRENCH, E., AND J. SONG (2013): "The Effect of Disability Insurance Receipt on Labor Supply, Appeals and Re-Applications: Evidence from Hearing Office Level Decisions," *American Economic Journal: Applied*, forthcoming.

GOTTSCHALK, P. (1996): "Is the Correlation in Welfare Participation Across Generations Spurious?," *Journal of Public Economics*, 63(1), 1 – 25.

GRUBER, J. (2000): "Disability Insurance Benefits and Labor Supply," *Journal of Political Economy*, 108(6), 1162–1183.

IMBENS, G. W., AND D. B. RUBIN (1997): "Estimating Outcome Distributions for Compliers in Instrumental Variables Models," *The Review of Economic Studies*, 64(4), 555–574.

KLING, J. R. (2006): "Incarceration Length, Employment, and Earnings," *The American Economic Review*, 96(3), 863–876.

Kostøl, A. R., and M. Mogstad (2013): "How Financial Incentives Induce Disability Insurance Recipients to Return to Work," *American Economic Review*, forthcoming.

Lee, C., and G. Solon (2009): "Trends in Intergenerational Income Mobility," *The Review of Economics and Statistics*, 91, 766–772.

Levine, P., and D. Zimmerman (1996): "The Intergenerational Correlation in AFDC Participation: Welfare Trap or Poverty Trap?," Discussion Paper 1100-96, Institute for Research on Poverty.

Maestas, N., K. Mullen, and A. Strand (2013): "Does Disability Insurance Discourage Work?," *American Economic Review*, forthcoming.

Mazumder, B. (2005): "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data," *The Review of Economics and Statistics*, 87(2), 235–255.

Milligan, K. S., and D. A. Wise (2011): "Social Security and Retirement around the World: Mortality, Health, Employment and Disability Insurance Participation and Reforms," *Harvard Kennedy School Working paper*.

Moffitt, R. (1992): "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature*, 30(1), 1–61.

Moore, T. J. (2011): "The Employment Effects of Terminating Disability Benefits: Insights from Removing Drug Addictions as Disabling Conditions," *Working Paper University of Maryland*.

OASDI-Trustees-Report (2012): "The 2012 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds," Discussion paper.

OECD (2010): *Sickness, Disability and Work: Breaking the Barriers. A Synthesis of Findings across OECD Countries*. OECD Publishing.

Oreopoulos, P., M. Page, and A. H. Stevens (2006): "Does Human Capital Transfer from Parent to Child? The Intergenerational Effects of Compulsory Schooling," *Journal of Labor Economics*, 24, 729–760.

Page, M. (2004): "New Evidence on the Intergenerational Correlation in Welfare Participation," chap. 10, pp. 226–244. Generational Income Mobility in North America and Europe, Cambridge University Press.

Page, M. E., and A. H. Stevens (2002): "Will You Miss Me When I Am Gone? The Economic Consequences of Absent Parents," NBER Working Papers 8786, National Bureau of Economic Research, Inc.

Parsons, D. O. (1991): "The Health and Earnings of Rejected Disability Insurance Applicants: Comment," *The American Economic Review*, 81(5), 1419–1426.

Pepper, J. (2000): "The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis," *The Review of Economics and Statistics*, 82, 472–288.

Rege, M., K. Telle, and M. Votruba (2009): "The Effect of Plant Downsizing On Disability Pension Utilization," *Journal of the European Economic Association*, 7(4), 754–785.

Solon, G., M. Corcoran, R. Gordon, and D. Laren (1988): "Sibling and Intergenerational Correlations in Welfare Program Participation," *The Journal of Human Resources*, 23(3), 388–396.

von Wachter, T., J. Song, and J. Manchester (2011): "Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program," *The American Economic Review*, 101(7), 3308–29.

Appendix Tables and Graphs

Table A.1: Descriptive Statistics.

Characteristic	DI applicants Mean	DI applicants Std. Dev.	DI appellants Mean	DI appellants Std. Dev.
Panel A: Parents				
Age (time of decision)	49.1	[4.6]	49.2	[4.4]
Female	0.658	[0.475]	0.735	[0.441]
Married	0.644	[0.479]	0.685	[0.465]
Foreign born	0.084	[0.277]	0.178	[0.382]
High school degree	0.445	[0.497]	0.374	[0.484]
College attendance	0.118	[0.323]	0.073	[0.260]
Children below age 18	0.422	[0.494]	0.427	[0.495]
Previous earnings ($), 1-10 years prior to decision	30,559	[22,263]	18,458	[19,179]
Years of work, 1-10 years prior to decision	8.0	[3.1]	6.0	[4.0]
Mental disorders	N/A	N/A	0.305	[0.460]
Musculoskeletal disorders	N/A	N/A	0.397	[0.489]
DI allowed	0.715	[0.451]	0.113	[0.317]
Number of parents	97,623		7,413	
Panel B: Children				
Age (time of decision)	25.4	[4.6]	25.0	[4.6]
Female	0.463	[0.499]	0.487	[0.500]
Married	0.156	[0.363]	0.164	[0.371]
Foreign born	0.097	[0.296]	0.127	[0.333]
High school degree	0.403	[0.491]	0.367	[0.482]
College attendance	0.172	[0.377]	0.117	[0.321]
Children below age 18	0.358	[0.479]	0.31	[0.463]
Previous earnings ($), 1-5 years prior to decision	18,322	[20,034]	20,716	[20,643]
Years of work, 1-5 years prior to decision	3.3	[1.9]	3.804	[1.633]
DI recipient 5 years after decision	0.038	[0.191]	0.027	[0.161]
DI recipient any time after decision	0.065	[0.247]	0.076	[0.266]
Number of children	200,866		14,893	

Notes: Sample of parents and children for applicants during the period 1992-2005 and appellants during the period 1989-2005. In both samples parents are restricted to be at most age 55 and their children to be aged 18 and above at the time of decision (at the application step or the appeal step). Previous earnings and years of work are measured the year before appeal in the DI appellant sample and the year before decision in the DI applicant sample. Nominal values are deflated to 2005 and represented in US dollars using the average exchange rate NOK/$ = 6. Unless otherwise stated, all parent and child characteristics are measured the year before parental application/appeal.

Table A.2: Characteristics of DI recipients in Norway and the U.S.

Characteristic	Norway	U.S.
Difficult to verify disorder	59.2%	57.3%
Age (at decision on initial application)	52.2	49.1
Prior earnings relative to the median	71.0%	69.9%

Notes: The U.S. numbers come from Maestas et. al (2012), and the Norwegian numbers are drawn from the sample of DI applicants during the years 2000-2003. Difficult to verify disorder include musculoskeletal and mental diagnoses. Prior earnings are measured 3-5 years before the application/appeal.

Table A.3: Characteristics of DI Applicants and Appellants in Norway and the U.S.

	Norway		U.S.	
Characteristic	Applicants	Appellants	Applicants	Appellants
Difficult to verify disorder	60.9%	69.7%	58.5%	62.2%
Age (at decision on initial application)	51.1	47.1	47.1	46.1
Prior earnings relative to the median	66.5%	50.4%	60.5%	56.3%

Notes: The U.S. numbers are from Maestas et. al (2012), and the Norwegian numbers are drawn from the sample of DI applicants during the years 2000-2003. Difficult to verify disorder includes musculoskeletal and mental diagnoses. Prior earnings are measured 3-5 years before the application/appeal.

Table A.4: Intergenerational Welfare Transmission using DI Applications as the Outcome.

	First stage		Reduced form		IV		N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
***Panel A:** Child applied for DI five years after parent's appeal decision*							
Parent allowed DI	0.913***	0.868***	0.057***	0.054**	0.063***	0.062***	14,893
	(0.113)	(0.115)	(0.021)	(0.021)	(0.022)	(0.023)	
Additional controls?	NO	YES	NO	YES	NO	YES	
Dependent mean	0.11	0.11	0.03	0.03	0.03	0.03	
***Panel B:** Child ever applied for DI after parent's appeal decision*							
Parent allowed DI	0.913***	0.868***	0.116***	0.110***	0.127***	0.127***	14,893
	(0.113)	(0.115)	(0.034)	(0.031)	(0.036)	(0.035)	
Additional controls?	NO	YES	NO	YES	NO	YES	
Dependent mean	0.11	0.11	0.09	0.09	0.09	0.09	

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: Baseline sample (see Table 3). Specifications mirror those of Table 3, except the outcome is child DI applications instead of child DI receipt.

Table A.5: **Characteristics of Marginal Applicants.**

Parental characteristic	First stage	$P[X=x]$	$P[X=x \mid complier]$	$\frac{P[X=x \mid complier]}{P[X=x]}$
Low education	1.028***	0.56	0.65	1.19
	(0.137)			
High education	0.704***	0.45	0.36	0.81
	(0.135)			
Young	0.870***	0.54	0.54	1.00
	(0.135)			
Old	0.876***	0.46	0.46	1.01
	(0.158)			
Married	0.919***	0.68	0.72	1.06
	(0.122)			
Not married	0.768***	0.32	0.28	0.89
	(0.165)			
High labor market experience	0.976***	0.43	0.48	1.12
	(0.185)			
Low labor market experience	0.806***	0.57	0.52	0.93
	(0.111)			

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

Notes: This table displays the first stage, marginal distribution, complier distribution and relative likelihood for different subgroups. Sample is restricted to appeals in the period 1989-2005. The first stage is estimated including the controls described in Table 3. Low education is defined as having 10 or fewer years of education, young as age 50 or less, and high labor market experience as working at least 9 out of the 10 years prior to the appeal decision.

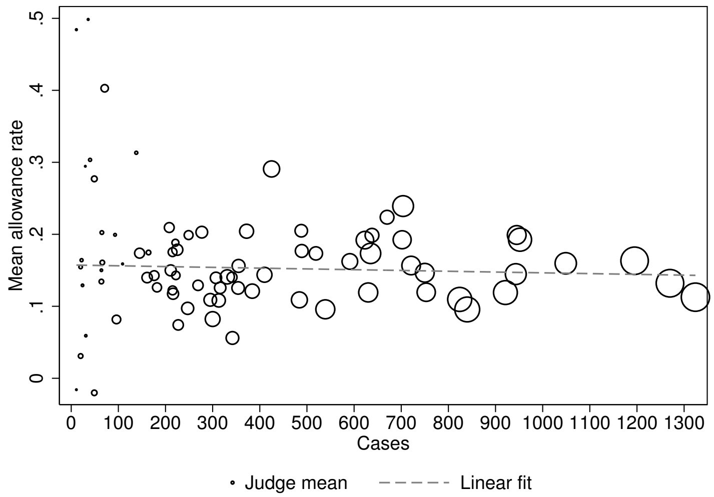
Table A.6: Effect of Judge Allowance on Child DI Participation: Re-weighted Estimates.

	Years since court decision				
	2 years	4 years	6 years	8 years	10 years
	Panel A: Baseline				
Estimate	0.038***	0.071***	0.109***	0.113***	0.125***
Standard Error	(0.011)	(0.017)	(0.025)	(0.027)	(0.036)
Dependent Mean	0.010	0.018	0.030	0.046	0.063
	Panel B: Re-weighted				
Estimate	0.034**	0.074***	0.115***	0.128***	0.127***
Standard Error	(0.014)	(0.020)	(0.028)	(0.028)	(0.035)
Dependent Mean	0.010	0.017	0.029	0.044	0.060

***p<.01, **p<.05, *p<.10. Standard errors (in parentheses) are clustered at the judge level.

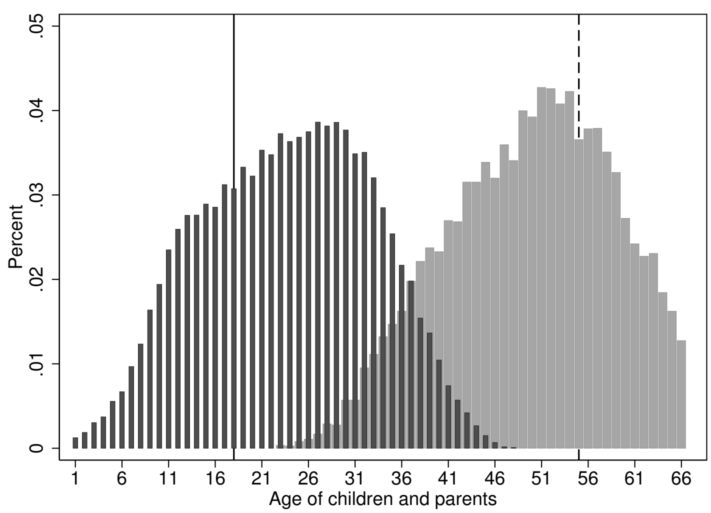
Notes: Balanced 10 year sample (see Table 6). For comparison, Panel A presents estimates from the top panel of Table 6. In Panel B these regressions are re-estimated after re-weighting individual observations so that the age distribution in each year is kept constant and centered around a mean age of 30.

Figure A.1: Judge Leniency versus Number of Cases Handled.



Notes: The figure plots a judge's allowance rate against the total number of cases he or she has handled. There are 79 different judges, and on average, each judge has handled a total of 380 cases. Allowance rates normalized by subtracting off year × department deviations from the overall mean. The sample is restricted to individuals appealing their first denied case during the period 1989-2005.

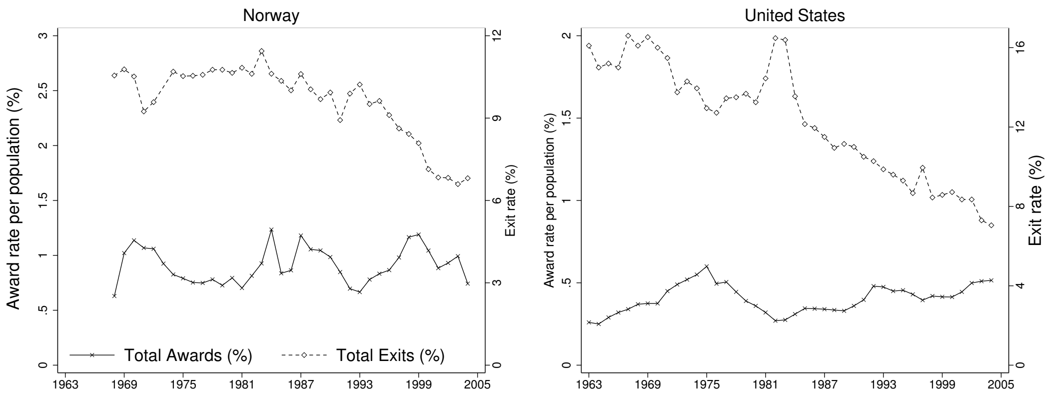
Figure A.2: Age Distribution of Parents and Children.



Notes: Age distribution for children (black) and parents (gray) meeting our other, non-aged based, sample restrictions. The baseline estimation sample is restricted to parents who are younger than 55 at the time of their appeal (denoted by the dashed vertical line) with children who are 18 or older (denoted by the solid vertical line).

Figure A.3:

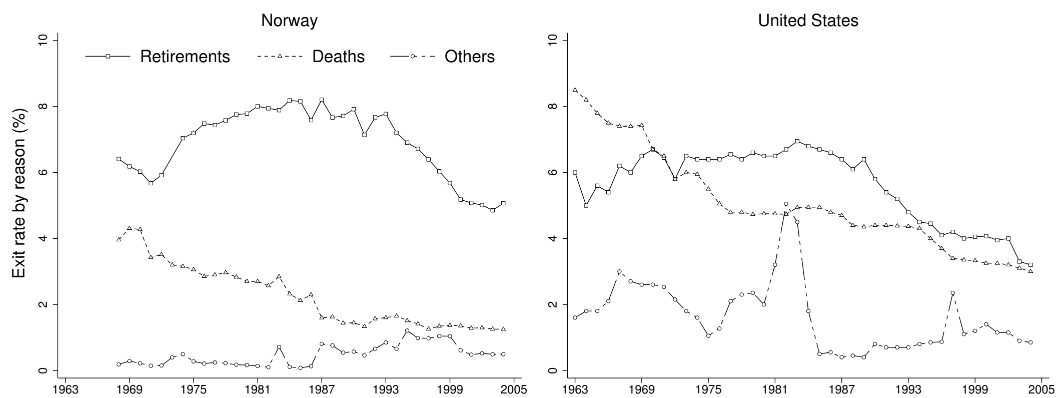
Award and Exit Rates



Notes: The U.S. trends are based on Autor and Duggan (2006), while the Norwegian trends are collected from various issues of the SSA Supplement. The graphs show award rates in the insured population and exit rates from the DI program in both countries.

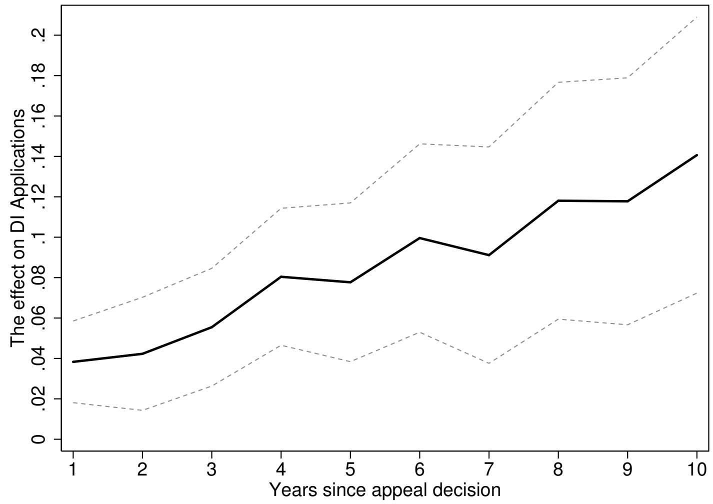
Figure A.4:

Exit rates by reason



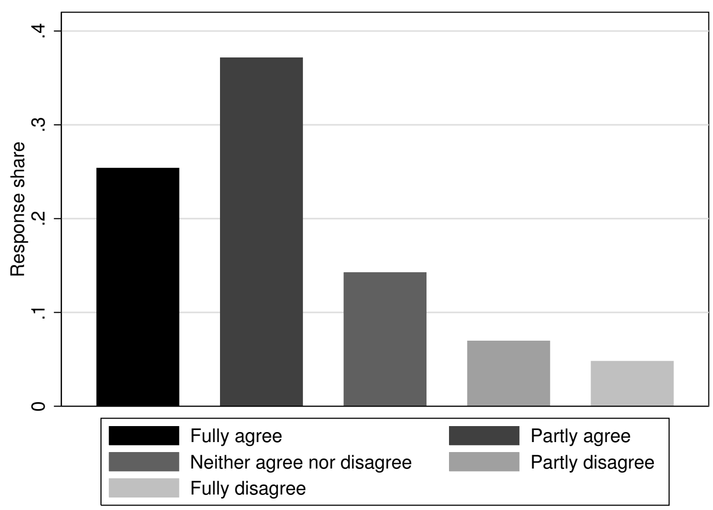
Notes: The U.S. trends are based on Autor and Duggan (2006), while the Norwegian trends are collected from various issues of the SSA Supplement. The graphs show exit rates because of death, retirement or other reasons (including eligibility-based exits).

Figure A.5: Intergenerational Transmission over Time using DI Application as the Outcome.



Note: This figure mirrors that of Figure 4 with the outcome being child DI application instead of child DI receipt.

Figure A.6: Survey Evidence on Social Assistance and Stigma.



Notes: Responses to the statement "Receiving social assistance makes people feel like second rate citizens" from a random sample of 3,190 Norwegians in 2007. Responses of "Don't know" (approximately 11%) are omitted from the figure.