Family Welfare Cultures

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

Preliminary Version: May 6, 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 less 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 because of a lenient judge, 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 rate 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.

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 more information about welfare eligibility than job availability; parents on welfare may provide negative peer pressure rather than positive role models; and parents on welfare may invest less 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. The literature has found strong correlations between parents' and children's participation in welfare, but has generally stopped short of establishing causality. Black and Devereux (2011), in their Handbook of Labor Economics chapter, summarize the state of the existing literature well: "while the intergenerational correlations in welfare receipt are clear, there is much less evidence that a causal relationship exists."

In this project, we investigate the existence and importance of family welfare cultures. We overcome the empirical challenges by exploiting a policy which randomizes the probability that parents receive welfare in combination with a unique source of population panel data that allows us to follow welfare recipients over time and match parents to children. We estimate the causal relationship in welfare participation across

[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 (Black and Devereux, 2011, Lee and Solon, 2009, Oreopoulos et al., 2006, Black et al., 2005), highlighting the difficulty of separating out correlations within families from causal effects. Levine and Zimmerman (1996) show that 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 credible causal inferences about intergenerational welfare transmission from observational data.

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.[3] Since its inception in 1961, DI rolls in Norway have grown from 2% to 10% of the adult population. These trends are not specific to Norway. For example, over the past 50 years DI rolls have steadily risen from less than 1% to over 5% in the U.S., and from 1% to 7% in the U.K. Many have argued that 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 that their adult children subsequently apply for and are awarded DI. Our approach takes advantage of the fact that the appeal judges are randomly assigned and therefore unrelated to any other factors (such as poverty or health) which might influence DI participation of the children. Our experimental research design can be interpreted as an instrumental variables (IV) model. The first stage uses the random assignment of appeal judges to obtain exogenous variation in DI participation of parents. The second stage examines whether this exogenous variation in DI participation of parents affects the probability that their children subsequently receive DI. To assess the internal validity of our research design, we perform a number of robustness checks, all of which suggest the IV 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 child's participation rate reaches 12 percentage points ten years after the judge's decision. We also explore the counterfactual labor market outcomes for children, i.e., what would have happened if a parent had not been allowed DI and therefore their child had not participated in DI. We

[3] In 2007, DI payments constituted 1.7% of GDP in the U.S. and 2.3% of GDP across the European OECD-countries (OECD, 2010). In 2011 the U.S. paid out $129 billion to 10.6 million disabled workers and their families, with an aditional $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 $9.6 billion to 4.6 million participants and SNAP (food stamps) paid out $80 billion to 46.5 million participants.

estimate that DI receipt decreases the probability that a child will be employed or attend college, reducing their labor force attachment and investments in human capital.

To better understand the IV estimates, we use the methods of Imbens and Rubin (1997) and Abadie (2003) to examine the type of 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. We show that merely one percent of the complier children would have been on DI if their parents had been denied DI. Although compliers only make up about a quarter of our sample, they are particularly relevant for policy, since reforms aimed at stemming the rise in DI will likely have the largest effect on these marginal applicants. 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 with difficult-to-verify disorders such as mental illness and musculoskeletal disorders (Autor and Duggan, 2006, Kostøl and Mogstad, 2013).

We further use the rich Norwegian data to explore the mechanisms behind the causal relationship in welfare participation across generations. Since we look at children who are at least 18 years old at the time of the parent's appeal decision,[4] 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 children as adults. We show, however, that the causal intergenerational relationship remains strong even if we exclude children who live at home with their parents or focus on children who are least 25 years of age and tend to have completed most of their schooling. 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 welfare use by a child do not go up after a parent is allowed DI, in contrast to what a model of general social stigma would suggest. Second, the estimated effect of a parent's DI experience on a child's DI participation increases over time. At the same time, 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 experience is consistent with these two pieces of evidence. In such a model, the children observe and respond to the initial rejection of parent's DI application and the subsequent loss in income; as initially denied parents begin to re-apply for DI, children learn the process is time consuming, risky and costly, which could dissuade them further from applying to DI.

[4] This age restriction is because individuals under 18 are not eligible for DI. While it would be interesting to study intergenerational welfare transmission among children who live at home and grow up with a parent on DI, we do not have a long enough panel to perform this analysis with precision.

Taken together, our results have important implications for the literature on intergenerational welfare transmission and for the evaluation of welfare programs. Our study provides novel evidence of welfare use in one generation causing welfare use in the next generation. This intergenerational link amplifies the direct effect of a change in a welfare program's generosity on parents, since their children's take up will also be affected. This leads to a long-run equilibrium participation rate which is substantially higher than otherwise expected. We simulate the total reduction in DI participation from a policy which makes the screening process more stringent. There is a direct effect on parents at the margin of program entry, but also an indirect effect operating through the decreased participation of their children. 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. However, over time, the direct effect shrinks, in part because some initially rejected parents re-apply and are awarded DI and in part because some parents reach retirement age and exit DI. 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. 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.

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.[5] Yet despite a recent surge in research on 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.[6] 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. Moreover, 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

[5]See e.g. Autor and Duggan (2003), Borghans et al. (2012), Bound (1989), Campolieti and Riddell (2012), French and Song (2013), Gruber (2000), Kostøl and Mogstad (2013), Maestas et al. (2013), Parsons (1991), Moore (2011), Von Wachter et al. (2011).

[6]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 et al. (2002), Bratberg (1999), Campolieti (2004), Gruber (2000), and Rege et al. (2009). None of these studies consider the role played by intergenerational welfare transmission.

affects labor supply.[7] Maestas et al. (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 the 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, the previous literature, 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, reports robustness checks, and examines heterogeneity in the estimated effects. Section 5 explores possible mechanisms and Section 6 presents a policy simulation. 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 et al. (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. 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:

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

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

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):

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

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, Levine and Zimmerman, 1996). 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 estimates are high. One needs 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 demonstrate the strong 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 the 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 variables.

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
Obs.	1,022,507	1,022,507	1,022,507
Dep. mean	0.03	0.03	0.03

***p<.01, **p<.05, *p<.10.

Notes: Standard errors are clustered at the family level. Data come from 2008 and are restricted to children age 23 and older with parents age 60 or younger and a grandparent who was alive during the period 1967-2010. 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). Column (3) controls flexibily for child, parent and grandparent characteristics (age, gender, education, foreign born, marital status, earnings history, and region fixed effects).

A large number of studies have used observational data to estimate models like equation (1).[8] As we do, they find strong intergenerational correlations. While these studies have helped researchers and policymakers 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 a causal intergenerational link.[9] 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

[8]See e.g. Antel (1992), Duncan, Hill, and Hoffman (1988), Gottschalk (1990; 1992; 1996), Moffit (1992), Page (2004) and Solon, Corcoran and Gordon (1988).
[9]See Gottschalk (1996), Levine and Zimmerman (1996), Pepper (2000), Bratberg et al. (2012) and Beaulieu (2005).

strong assumptions or using standard instruments, he finds the bounds are wide and the point estimates are noisy and often inconsistent across specifications. His estimates suggest, however, that OLS suffers from omitted variables bias.

A solution to the problems encountered in the literature would be to randomly assign parents to a treatment group which participates in a welfare program and a control group which does not. In this setting, there would be no omitted variable bias, since a parent's earnings potential and health (and all other characteristics) would be uncorrelated with program participation. This is the idea of our research design, which we discuss in the next subsection.

2.2 Experimental Setting and Research Design

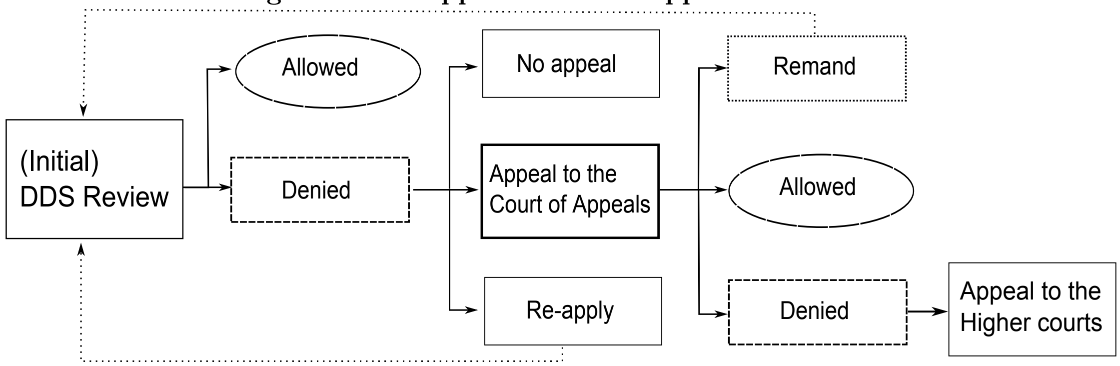
In this subsection, we begin by reviewing key facts regarding the DI program in Norway. We then provide institutional details and empirical evidence on the disability determination process, documenting in particular that the system generates random variation in DI awards. We further describe how we will 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 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 make one of three decisions: allowed, denied, or remand.[10] 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 then appeal to the higher court level. Very few individuals go through this final stage of the disability determination process.

Random Assignment of DI Cases to Judges.

In Norway, the hearing of appeals is centralized in Oslo, handling all cases for the entire country. Prior to 1998, there was only one department. Afterwards, there were four equally-sized departments; 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 occurs by random allocation, as stipulated in the rules set forth for the Administrative Law Court since its inception in 1967. The rules state that assignment should be done "by the drawing of lots." In practice, cases are assigned randomly 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 any other aspect of the case, (iii) the

[10] A remand is a case which the judge sends back to the disability determination stage with a request for more information. In our baseline analysis, we code a remanded case as rejected. In a robustness check, we code remanded cases as allowed or denied based on its eventual outcome after it is reconsidered by the judge with updated information.

[11] We verified this with Knut Brofoss, the current Head of the Administrative Law Court.

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 the other cases a judge has handled. 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.[12]

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. Figure 2 plots 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.

[12]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: Test for Random Assignment of Judges: OLS Estimates of Applicant Characteristics on Case Allowance and Judge Leniency.

	Dependent Variable			
	Case Allowed		Judge Leniency	
	coeff.	s.e.	coeff.	s.e.
Age	0.00550***	(0.00092)	0.00037*	(0.00019)
Female	0.01411	(0.00957)	0.00047	(0.00184)
Married	0.00677	(0.00748)	0.0015	(0.00190)
Foreign born	-0.02779***	(0.01166)	0.00115	(0.00241)
High school degree	0.01256*	(0.00728)	0.00035	(0.00141)
Some college	0.02587	(0.01705)	-0.00057	(0.00330)
College graduate	-0.09746***	(0.01776)	0.00417	(0.00947)
One child	-0.0064	(0.00880)	-0.00113	(0.00197)
Two children	-0.01353	(0.01354)	0.00112	(0.00158)
Three or more children	-0.03437***	(0.01406)	0.00262	(0.00220)
Average indexed earnings	0.00000***	(0.00000)	0.00000	(0.00000)
Experience	0.00730***	(0.00087)	0.00007	(0.00019)
Mental disorders	0.02815*	(0.01470)	0.00136	(0.00598)
Musculoskeletal disorders	0.06511***	(0.01630)	0.00077	(0.00593)
F-test for joint significance	9.9627		.7367	
[p-value]	[.001]		[.7314]	

***p<.01, **p<.05, *p<.10.

Notes: Standard errors are clustered at the family level. There are 14 894 individual observations and 79 different judges. Sample restricted to individuals appealing their first denied case during the period 1989-2005.

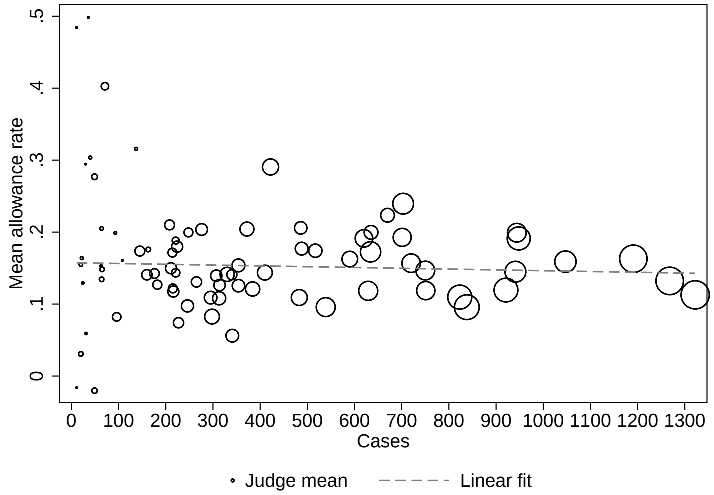
Using Judge Leniency as an Instrument.

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 children's 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) health, earning capacity, 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:[13]

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

[13]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.

Figure 2: 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. Allowance rates normalized to subtract off year × department deviations from the overall mean.

Although we do not observe a judge's leniency directly, we can consistently estimate it by taking the average allowance rate in the other cases he or she has handled, as we did for Table 2.[14] 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, where equation (4) is the first stage and equation (1) is the second stage in a two-stage least squares regression. 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[15], the month when DI was awarded (or denied), and

[14]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. 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.

[15]Unfortunately, the medical diagnoses are only availabe from 2000.

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 the unique identifiers of both the judges and the 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 (regarding sex, age, and number of children), socio-economic data (regarding years of education and earnings), and geographical 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 e.g. Atkinson et al. 1995).

Our empirical analysis considers children of parents who appeal an initially denied DI claim. Following Maestas et al. (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 restrict the sample to children who were at least 18 years old at the time of the parent's appeal decision. This age restriction is because individuals under 18 are not eligible for DI. Lastly, we 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 substitition between DI and early retirement schemes.

In Appendix Table A.1, we document the key characteristics of the sample of parents who apply for DI (Panel A) and those who appeal an initially denied DI claim (Panel B). The parents who appeal are on average more likely to be female, less educated, foreign born, and have less work experience as compared to the group of initial applicants. The children of parents who appeal tend to be less educated and more likely to eventually end up on DI as compared to the children of parents who apply for DI.

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.[16] 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 3 shows this distinction by displaying the evolution

[16]Our discussion of the U.S. system draws primarily on Autor and Duggan (2006), while our discussion of the Norwegian system is based on Kostøl and Mogstad (2013).

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.,[17] the time trends are quite similar. From 1961 to 2012, the rate of receipt increased from 2.2 to 10.4 percent in Norway and from 0.8 to 5.3 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).

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.[18] Because these are early-onset disorders with low age-specific mortality, DI recipients with such diagnoses tend to participate in the program for relatively long periods. As a result, the DI exit rates have decreased in the last few decades. Appendix Figure A.1 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.2, 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.

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.3 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.[19] 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.

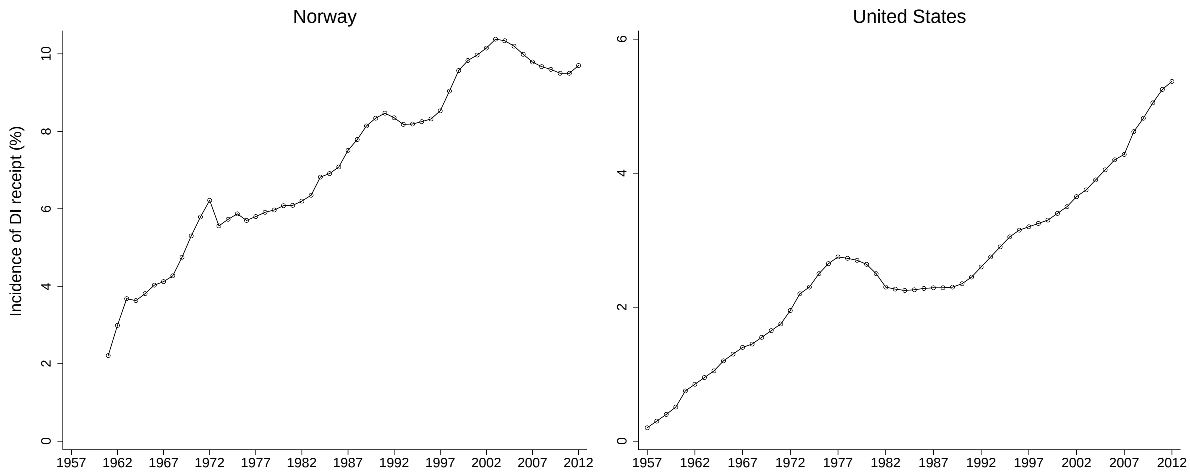
A third difference is that the appeal process plays a more important role in the U.S. than in Norway. In both countries, the disability determination process is a multi-step process, where cases that are difficult to judge are often denied at the initial application step. If the DI claim is initially denied, the individual

[17]The cross-country difference in coverage by the DI program 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 rates of DI receipt across developed countries.

[18]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.

[19]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 2011 in the U.S., direct expenditures on DI totalled \$129 billion and Medicare expenditures for disabled workers totalled \$90 billion. In Norway, all citizens are eligible for health insurance through the Social Insurance System.

Figure 3: 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 2006-1012. Norwegian trends based on SSA Statistical Supplements. Incidence of DI receipt defined as the percent of the relevant adult population receiving DI benefits (ages 18 - 67 in Norway; ages 25-64 in the US).

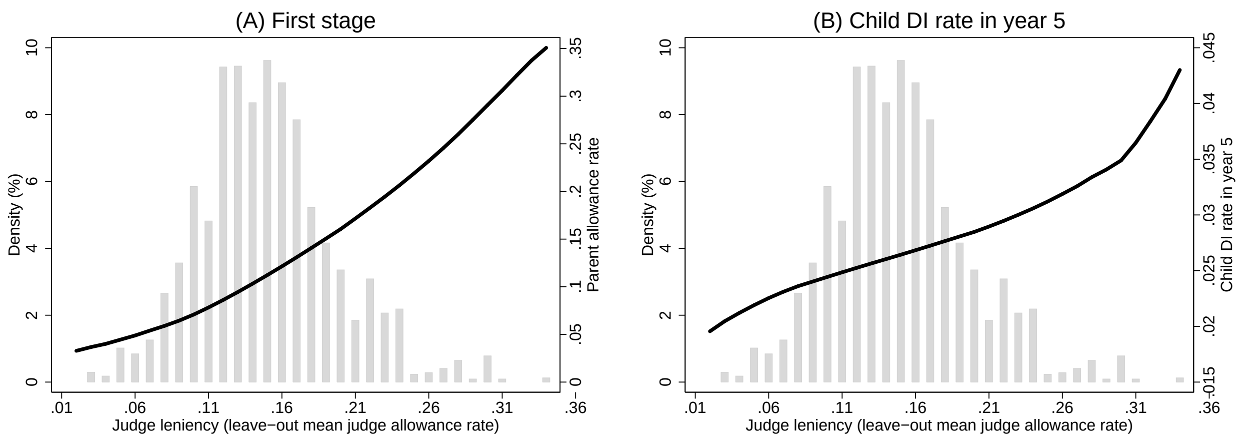
may appeal the decision to the Court of Appeals where the appeals are reviewed by Administrative Law Judges. While 48 percent of the initially rejected applicants appeal in the U.S. (French and Song, 2013), only 25 percent of the initially rejected appeal in Norway. Appendix Table A.4 compares the characteristics 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, the 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 4. In the background of each graph is a histogram for the density of judge leniency, which captures the average 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. 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 20% of cases allowed by a judge at the 90th percentile compared to approximately 6% at the 10th percentile.

Figure 4: Effect of Judge Leniency on Parents (First Stage) and Children (Reduced Form).



Notes: 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. 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 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 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 0.92 percentage points. This effect is not statistically different from one.

Table 3: Estimates of Intergenerational Welfare Transmission.

	First stage		Reduced form		2SLS		N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: *Child on DI five years after parent's appeal*							
Parent allowed DI	0.918***	0.873***	0.054**	0.051**	0.059***	0.059***	14,894
	(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*							
Parent allowed DI	0.918***	0.873***	0.108***	0.104***	0.117***	0.119***	14,894
	(0.113)	(0.115)	(0.029)	(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.

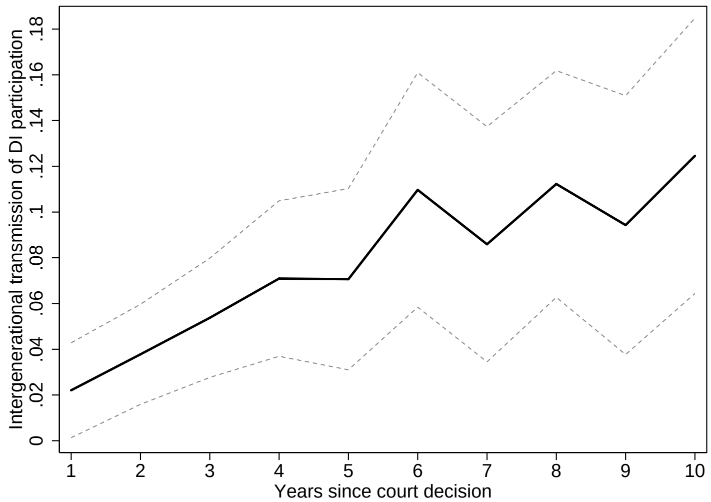
Notes: Standard errors clustered at the judge level.

Panel A shows results for the child, where the outcome variable is whether the child participates in DI within 5 years after the parent's initial appeal. 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 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 2SLS estimates are very similar.

Panel B performs a similar exercise, but now looks at whether the child has ever been on DI after their parent's appeal. While every child is observed for at least five years, in this second panel some children will be observed for up to 21 years after their parent's appeal. 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 that the long-run effects of a parent getting on to DI are larger than the short-run effects: the 2SLS 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 5 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. 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 5: Estimates of Intergenerational Transmission over Time.



Notes: Solid line is the yearly estimated effect of a parent's DI participation, instrumented with judge leniency, on their child's participation. Data restricted to pre-2000 appeals so as to have a balanced 10 year sample. Dashed lines represent 95 percent confidence intervals.

Lastly, we shift attention to how a parent's DI participation affects the probability that their children subsequenty apply for DI. Appendix Figure A.3 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 application grows substantially over time. Ten years after the court decision, the causal effect of a parent being allowed DI is a 14 percentage point increase in a child's DI application rate. Given the similarity in the estimates, we will focus on child DI participation in the remainder of the paper.

4.3 Internal Validity

In order for judge leniency to be a valid instrument, the 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 judges to individuals who appeal their cases. The even numbered columns of Table 3 explore what happens if a large set of control variables are added to the regressions. This can be viewed as a second test for the random assignment of judges. 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 randomizaton, 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 serious 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.[20]

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 additional assumption 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 causal effects of judge leniency. One attractive feature of the process in Norway makes it likely that the exclusion restriction holds: 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). If that is the case, and the appeal processing time affects a child's choice of applying for DI, then our instrument would operate through more than one causal channel. It is reassuring to find that our instrument, judge leniency, and judge processing time are virtually uncorrelated.[21] Moreover, the first row of Table 4 shows that the IV estimates do not change appreciably if we control for judge processing in the first and second stage.

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. A failure of monotonicity means a lower judge leniency value

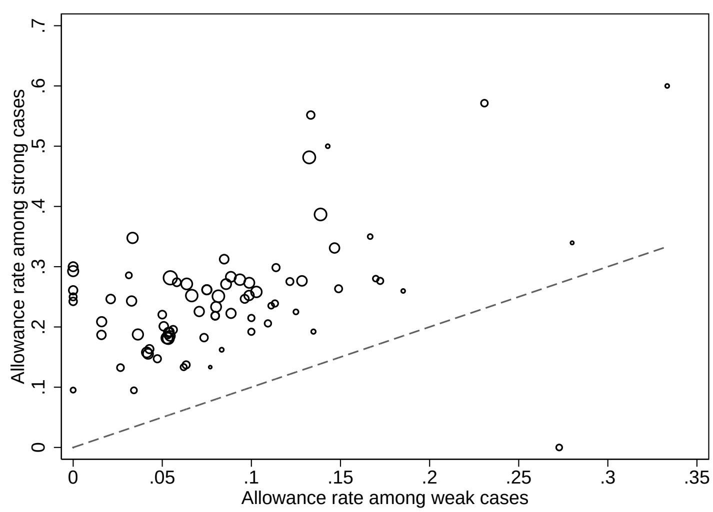
[20] The regression coefficient of judge leniency on child's partication in the sickness program is 0.013 (s.e. = 0.054). This point estimate is low compared to the sample mean: 23 percent of the children had received sickness pay at some point prior to the parent's appeal.

[21] To explore this possibility, we calculated judge processing time based on the residual average processing time in the other cases a judge has handled after controlling for a fully interacted set of time and department dummies in a regression. The regression coefficient of judge processing time on judge leniency is -0.001 (s.e. = 0.003).

pushes some parents onto DI while a higher value pushes others out, which would bias the IV estimates if the effects of parental DI participation in the former group differ from the effects in the latter group. 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 to one another for similar cases. To conduct this test, we use the regression estimates from column 1 in Table 2 to predict the probability a case will be allowed. Based on these predictions, we create two equally-sized groups: appellants who have a strong case (an above median probability of allowance) and those who have a weak case (a below median allowance probability). For each judge, 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 judges in our sample, as graphed in Figure 6. That figure plots each judge's allowance rate in weak cases versus their allowance rate in strong cases, with all observations lying above the 45 degree line.[22]

Figure 6: Judge Allowance Rates in Strong vs. Weak Cases



Note: Each dot plots an individual judge's average allowance rate in their stronger versus weaker cases. A strong (weak) case is defined as having a predicted allowance probability above (below) the median based on estimates from column 1 of Table 2. Dots size is proportional to the t-statistic for the difference in means. Dashed line represents the 45 degree line.

[22]While the grouping into strong and weak cases is based on a regression using all cases (including the appellant's own case), this should not overly influence each judge's allowance rates since the number of cases is reasonably large. On average, the judges in our sample handled 690 cases.

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 that the results are robust to adding in month-department controls or excluding parents who die. In our baseline analysis, we excluded judges who handle less than 10 cases. The fifth and sixth row 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. If we instead code remanded cases as allowed or denied based on its eventual outcome after it is reconsidered by the judge with updated information, the results are quite similar.

Table 4: Specification Checks for Child on DI Five Years after Parent's Appeal

	First stage	IV	N
With judge processing time	0.860***	0.059***	14,894
	(0.107)	(0.022)	
One Department (pre-1998)	1.101***	0.054**	5,590
	(0.187)	(0.025)	
Month-department controls	0.779***	0.065**	14,894
	(0.128)	(0.029)	
Exclude parents who die	0.864***	0.066***	14,505
	(0.113)	(0.023)	
Include judges < 10 cases	0.860***	0.058**	14,898
	(0.116)	(0.023)	
Exclude judges < 50 cases	0.949***	0.055**	14,759
	(0.110)	(0.021)	
Alternative coding of remand	0.804***	0.064***	14,894
	(0.103)	(0.022)	

***p<.01, **p<.05, *p<.10.

Notes: Standard errors are clustered at the judge level. The specifications mirror those of Table 3.

4.4 Interpreting IV estimates

Our IV estimates should be interpreted as a local average treatement effects (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 our instrument of 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.[23] 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 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 11.9 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.[24] We estimate that only 1.2 percent of the complier children would have 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 13.1 percent. Lastly, we characterize marginal applicants by observable characteristics in Appendix Table A.5. The most distinctive feature of the marginal applicants is their educational attainment: Sixty-six percent of the compliers have low education, while their fraction of the entire sample is only 55 percent. This finding indicates that a tightening of the screening process would disproportionately affect low educated parents and their children.

[23]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.

[24]Let $P^c(0)$ and $P^c(1)$ denote potential outcomes for children based on whether or not their parent is allowed DI, respectively. Continuing with the notation developed in footnote 23, $E(P^c(0)|n) = E(P^c|P^p = 0, z_i^p = \overline{z})$ and $E(P^c(1)|a) = E(P^c|P^p = 1, z_i^p = \underline{z})$. The potential outcomes of compliers 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 $E(P^c \mid P^p = 1, z_i^p = \overline{z}) = \frac{\pi_c}{\pi_c + \pi_a} E(P^c(1) \mid c) + \frac{\pi_n}{\pi_c + \pi_a} E(P^c(1) \mid a)$. The LATE is given by $E(P^c(1)|c) - E(P^c(0)|c)$.

4.5 Heterogeneity by Network Characteristics

In Table 5, we look at subsample estimates based on network characteristics. We first split the sample into neighborhoods which have below median and above median DI participation rates (the median is 8.2 percent). We define a neighborhood as a street. The intergenerational transmission estimate is larger for neighborhoods with a high DI rate Specification 2 in the table performs a similar analysis, but this time splits the sample into neighborhoods with below median and above median employment rates (the median is 71.4 percent). There is a hint that low employment neighborhoods have larger estimates, but the estimates are quite imprecise.

In the final specification of Table 5, we look at extended family networks. We consider subsamples where no extended family member is currently on DI versus where at least one extended family member is currently on DI. The extended family network includes aunts, uncles, siblings, and cousins. Here the subsamples yield different point estimates: children who have no extended family members on DI are more influenced by a parent's participation in DI, although the difference is not statistically significant. The differences could be driven by a larger reduction in stigma when a parent is the first person in the extended family to go on to DI, or by a relatively larger information gain. We explore which of these two explanations is more likely in Section 6.

Table 5: Subsample Estimates for Child on DI Five Years after Parent's Appeal

	First stage	IV	N
1. Low DI rate	0.864***	0.042	7,455
in neighborhood	(0.123)	(0.027)	
High DI rate	0.872***	0.088**	7,439
in neighborhood	(0.121)	(0.039)	
2. High employment rate	0.937***	0.056*	7,447
in neighborhood	(0.141)	(0.032)	
Low employment rate	0.786***	0.071**	7,447
in neighborhood	(0.122)	(0.036)	
3. No one on DI	0.874***	0.075**	9,214
in extended family	(0.127)	(0.031)	
At least one on DI	0.849***	0.048	5,680
in extended family	(0.152)	(0.047)	

***p<.01, **p<.05, *p<.10.

Notes: Standard errors are clustered at the judge level. All networks are defined one year before the parent appeal. Neighbourhood networks are divided into low (high) participation rate if the network average is below (above) the overall median. Median DI rate is 8.2 percent, while median employment rate is 71 percent..

5 Policy Simulation

Our results provide strong evidence that welfare use in one generation causes welfare use in the next generation. These intergenerational effects are important for policy, as they indicate that any changes to the DI system will affect not only current applicants, but also have important spillover effects on their children. The intergenerational link will amplify the direct effect of a change in a welfare program's generosity on parents, since their children's take up will also be affected. This leads to a long-run equilibrium participation rate which could be substantially higher or lower than otherwise expected.

In this section, we simulate the total reduction in DI participation from a policy which makes the screening process more stringent. 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 could be acheived 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. There are two components to the total reduction in DI from this 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 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 multiplying this estimated coefficient by 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. As in Figure 5, we restrict the sample to children observed for 10 years after their parent's appeal decision. The effect of judge leniency on child DI participation grows substantially 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 the retirement age and exit DI.

Using the estimates in Table 6, Figure 7 graphs the results of the policy simulation over time. In the first year after the judge's 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

Table 6: Effect of Judge Allowance on Child and Parental DI participation

	Year 2	Year 4	Year 6	Year 8	Year 10
			Panel A: DI Child		
Estimate	0.038***	0.071***	0.110***	0.112***	0.125***
St. Error	(0.011)	(0.017)	(0.026)	(0.025)	(0.031)
Dep. Mean	0.01	0.018	0.03	0.046	0.063
			Panel B: DI Parent		
Estimate	0.700***	0.642***	0.440***	0.390***	0.273**
St. Error	(0.113)	(0.142)	(0.133)	(0.112)	(0.117)
Dep. Mean	0.385	0.508	0.581	0.631	0.669

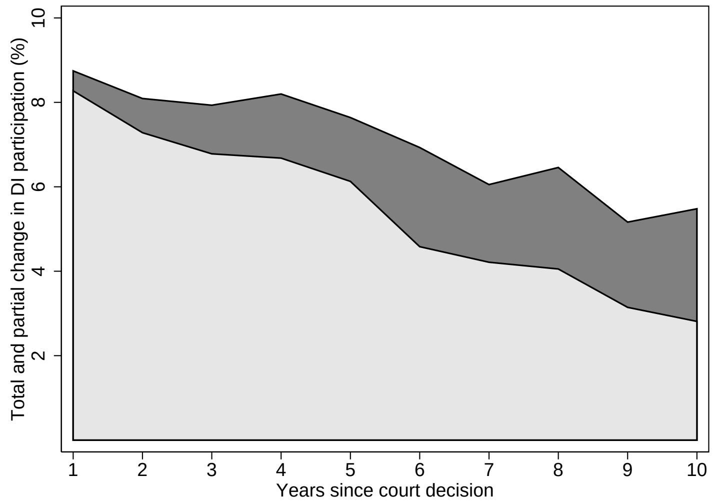
***p<.01, **p<.05, *p<.10.

Notes: Standard errors are clustered at the judge level. There are 9144 individual observations and 50 different judges in the balanced sample for the years 2, 4, 6, 8 and 10. The table presents IV-estimates on DI participation for parents and children for these years.

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.

This simulation makes clear that failing to account for intergenerational effects will provide misleading cost estimates. To translate the participation patterns shown in Figure 7 into cost terms, we calculated the net present value of the simulated policy change for parents and children over time, based on the average DI benefits amount and assuming a 5 percent annual discount rate. Making judges approximately 10 percent stricter (one-fifth of a standard deviation reduction in judge leniency) decreases the net present value of program expenditures after 10 year 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 half 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 retirment, while parents are older and age out of the system sooner.

Figure 7: The Effect of Tightening the Screening Process on Parents and Their Children.



Notes: Light grey area is the direct effect of tightening the screening process by one-fifth of a standard deviation (i.e., allowing approximately 10% fewer DI cases) on parents' participation. Dark grey area is the spillover effect on their children due to the intergenerational transmission of DI use.

6 Mechanisms

6.1 What is the Counterfactual for Parents and Children?

In this section, we estimate the effect of DI receipt on labor market outcomes, both for the parent and their child. This information is useful in understanding not only labor supply effects for parents and their children, 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.

To qualify for DI benefits, a worker must demonstrate that they have a physical or mental health impairment which substantially reduces their ability for gainful employment. Before going onto DI, a worker usually participates in the long-term sickness program; these benefits last for one year and provide 100% income replacement. Workers are also required, if deemed able, to participate in an employment rehabilitation program. While a worker can have limited labor income and still qualify for DI benefits, earnings are not allowed to exceed the "substantial gainful activity" (SGA) level once on DI. In 2005, the threshold was approximately $10,000 annually. Combined, these requirements mean that parents already have a limited attachment to the labor market by the time they apply for DI benefits.

Table 7 estimates how DI receipt affects labor outcomes, using our judge leniency measure as an instrument. In Panel A, which shows results for parents, DI receipt is associated with a sizeable drop in the fraction of parents receiving more earnings than the substantial gainful activity threshold. Mean labor income and the probability of full-time work also fall substantially for parents on 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 $7,977 and the probability of labor income being above the SGA threshold (full-time work) to decrease by 19 (11.9) percentage points.

Table 7: Effect of DI on Parental and Child Labor Outcomes Five Years after Parental Appeal.

	Reduced form	2SLS	Dep. mean	N
	Panel A: Parents			
1. Labor Income>SGA	-0.166**	-0.190**	0.31	14,894
	(0.072)	(0.079)		
2. Labor Income ($)	-6,960***	-7,977***	10,070	14,894
	(2,509)	(2,868)		
3. Full-time work	-0.104***	-0.119***	0.04	14,894
	(0.038)	(0.042)		
	Panel B: Children			
4. Employment	-0.134**	-0.154**	0.48	14,894
	(0.052)	(0.063)		
5. Full-time work	-0.094	-0.108	0.36	14,894
	(0.073)	(0.083)		
6. College attendance	-0.055	-0.063	0.23	14,894
	(0.060)	(0.068)		

***p<.01, **p<.05, *p<.10.

*Notes:*Standard errors are clustered at the judge level. The SGA level in 2005 equals 10, 000 USD. Full-time work is defined as working more than 30 hours a week. Employment is defined as working more than 4 hours a week. College atteendance is having completed a higher degree by the end of our observation window (2010).

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 et al. (2013) exploit the leniency of initial examiners in the 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 counterfactual for parents is comparable across the two countries.

Panel B in Table 7 explores the counterfactual labor outcomes for children, i.e., what would have happened if a parent had not been allowed on DI and therefore the child had not participated in DI? Examining child outcomes five years after their parent's appeal, we find that a child's DI receipt causes

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

	Five years after parent's appeal			
	Reduced form	2SLS	Dependent mean	N
1. Child living away from home	0.076**	0.080***	0.03	8,652
	(0.031)	(0.031)		
2. Child at least 25 years of age	0.075**	0.072**	0.03	6,562
	(0.029)	(0.029)		

***p<.01, **p<.05, *p<.10.

Notes: Standard errors clustered at the judge level. Child residency is based on whether having a different address than the parent one year before the parent appealed.

employment (at least 4 hours a week) to drop by more than 15 percentage points. While we do not have enough precision to estimate the drop in full-time work or college attendance precisely, both estimates suggest a sizeable drop in these child outcomes as well. The reduced labor force attachment for children is especially important from a policy standpoint. Since few individuals exit DI after entering, any reductions in employment or increases in DI payments will likely last for a large number of years until children reach retirement.

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 the welfare programs and less information about employment, and (iii) parents on welfare may invest less 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.

Since we look at children who are at least 18 years old at the time of the 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 children as adults. In Table 8, however, we show that 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 of age and tend to have completed most of their schooling. While it is possible that parents who go on DI provide income assistance to older children who do not live at home, and that this in turn causes them to participate in DI at a higher rate, we believe this is an unlikely mechanism.

Table 9: The Effect of Parental DI Allowance on Child Participation in Various Welfare Programs

	Five years after parent's appeal			
	Reduced form	2SLS	Dependent mean	N
1. Child on disability insurance	0.051**	0.059***	0.03	14,894
	(0.020)	(0.022)		
2. Child on social assistance (i.e., traditional welfare)	-0.008	-0.010	0.10	14,894
	(0.046)	(0.053)		
3. Child receiving other cash transfer	-0.002	-0.002	0.61	14,894
	(0.062)	(0.070)		

***p<.01, **p<.05, *p<.10.

Notes: Standard errors clustered at the judge level.

This 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 complimentarity in DI; if a parent is on DI, a child's utility of taking up DI is 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 assess the role of information transmission. However, we find two pieces of suggestive evidence against the hypothesis that our findings are driven by a drop in social stigma resulting from parental DI use. Our first piece of evidence considers DI in the broader context of all welfare programs, including traditional welfare (labelled "social assistance" in Norway)[25] and other cash transfer programs (a broad measure which captures unemployment insurance payments, benefits for single parents,child allowances, housing benefits, etc.). 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) and Aaberge and Flood (2012). In our setting, if a parent participates in DI, it is likely the stigma of participating in other welfare programs would also go down somewhat.

Table 9 reports estimates of the causal 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

[25]The Norwegian social assistance system differs somewhat from traditional welfare programs in other countries. It is considered a last-resort safety net. Benefits are means tested and provided by local governments. There are no clear rules about eligibility or benefit amounts, with discretion being left to the local social worker.

assistance program (traditional welfare) on their parent's DI allowance. Both the reduced form and the 2SLS 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 = 0.099). When we look at whether a parent's DI experience affects the likelihood a child receives any other type of cash transfer, we similarly find no effect.

Table 9 does 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 second test for stigma versus information relies on how DI take up rates for both parents and their children vary over time. The first thing to note is that many parents who are initially denied re-apply and are eventually allowed DI. Figure 8 shows the time profile of DI participation for parents who are initially allowed DI versus those who are initially denied at the appeals stage. The dashed line documents that over 95% of parents allowed DI actually take up DI. It takes a year or so for these allowed parents to transition to DI, but after that, the participation rate for allowed parents remains high with few parents exiting the program. In contrast, the fraction of initially denied parents receiving DI benefits rises monotonically over time, as plotted by the solid line. As expected, the year of the initial appeal there are few denied parents on DI, but 10 years later, over two-thirds of of these initially denied parents are receiving DI benefits. This time pattern makes sense as it takes some time to re-apply for DI and some individuals apply multiple times.

The time patterns in Figure 8 have different predictions for the stigma and learning channels. If 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. In contrast, social learning predicts that a child learns from their parent's experience. Since re-application is time consuming, risky and costly, as this information is revealed it should dissuade children from applying. Indeed, there are large income losses to the parent during re-application, because if an applicant demonstrates that they have substantial earnings capacity, they are less likely to be allowed DI.[26] The majority of applicants do not return to the labor market after a judge dismisses their case, and if they do, full-time work is rare.

Of course, it could also be argued that a child with rational expectations could predict the costs and

[26]Note that all parents in our sample make it to the DI appeal stage, so our estimate is not driven by information about how to initially apply to the DI program.

Figure 8: Fraction of Allowed and Denied Parents Receiving DI Benefits over Time.

benefits of re-application, in which case there would be no change in a child's DI participation over time. One could also argue 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 shrinks over time. This would be a case of positive information being revealed over time, rather than negative information.

In summary, if stigma or learning about positive information is the primary channel, Figure 8 predicts that the effects on children's participation should fade out over time as more denied parents get on DI. In contrast, if children are learning that re-application is both risky and costly, the gap in DI participation for treatment and control children should increase over time. The pattern in Figure 5 fits poorly with the hypothesis that stigma is responsible for the causal effects we estimate. Likewise, positive information or a story of unbiased forward looking expectations is not consistent with the data. By comparison, a model where children are learning that DI re-application is risky or costly is consistent with the observed time-series pattern.

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 some of the first quasi-experimental evidence that such a causal link exists. The key to our identification approach is that judges are randomly assigned to DI applicants whose cases are initially denied and some appeal judges are systematically more lenient. Using this random variation, we find strong evidence that 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 rich administrative data allow us to explore the mechanisms for the large causal effects we estimate. By looking at participation in other welfare programs besides DI and the time series pattern in take-up, we find evidence against the hypothesis that a drop in stigma resulting from parental DI use is the primary driver of our results. Our results also suggest that parental investment in children is an unlikely mechanism. Rather, our findings are consistent with a model where children learn from parents that applying and re-applying to the DI program is risky and costly in terms of both time and wages.

Our results have important implications for the literature on intergenerational welfare transmission and for the evaluation of welfare programs. Our study points out that welfare use in one generation causes welfare use in the next generation. This intergenerational link amplifies the direct effect of a change in a welfare program's generosity on parents, since their children's take up will also increase, and can lead to a long-run equilibrium participation rate which is substantially higher than otherwise expected. In terms of program expenditure, is 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.

References

Aaberge, R. and Flood, L. (2012). Us versus sweden: The effect of alternative in-work tax credit policies on labour supply of single mothers. *Working Paper*.

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2):231 – 263.

Atkinson, A., Rainwater, L., and Smeeding, T. (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*, No. 17697.

Autor, D. H. and Duggan, M. G. (2003). The rise in the disability rolls and the decline in unemployment. *Quarterly Journal of Economics*, 118(1):157–205.

Autor, D. H. and Duggan, M. G. (2006). The growth in the social security disability rolls: A fiscal crisis unfolding. *Journal of Economic Perspectives*, 20(3):71–96.

Bertrand, M., Luttmer, E. F. P., and Mullainathan, S. (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.

Black, D., Daniel, K., and Sanders, S. (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 Devereux, P. (2011). *Recent Developments in Intergenerational Mobility*, chapter 16, pages 1487–1541. Handbook of Labor Economics, Elsevier.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review*, 95(1):437–449.

Blundell, R. and Macurdy, T. (1999). Chapter 27 labor supply: A review of alternative approaches. volume 3, Part A of *Handbook of Labor Economics*, pages 1559 – 1695. Elsevier.

Borghans, L., Gielen, A., and Luttmer, E. (2012). Social support subbstitution 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., Nilsen, O. A., and Vaage, K. (2012). Is recipiency of disability pension hereditary?

Burkhauser, R. V. and Daly, M. C. (2012). Social security disability insurance: Time for fundamental change. *Journal of Policy Analysis and Management*, Forthcoming.

Campolieti, M. (2004). Disability insurance benefits and labor supply: Some additional evidence. *Journal of Labor Economics*, 22(4):863–889.

Campolieti, M. and Riddell, C. (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.

Doyle, J. (2008). Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care. *Journal of Political Economy*, 116(4):746–770.

French, E. and Song, J. (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 Rubin, D. B. (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 Mogstad, M. (2013). How financial incentives induce disability insurance recipients to return to work. *American Economic Review*, forthcoming.

Lee, C.-I. and Solon, G. (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics*, 91:766–772.

Levine, P. and Zimmerman, D. (1996). The intergenerational correlation in afdc participation: Welfare trap or poverty trap? Discussion Paper 1100-96, Institute for Research on Poverty.

Maestas, N., Mullen, K., and Strand, A. (2013). Does disability insurance discourage work? *American Economic Review*, forthcoming.

Milligan, K. S. and Wise, D. A. (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*.

OECD (2010). *Sickness, Disability and Work: Breaking the Barriers. A synthesis of findings across OECD countries*. OECD Publishing.

Oreopoulos, P., Page, M., and Stevens, A. H. (2006). Does human capital transfer from parent to child? the intergenerational effects of compulsory schooling. *Journal of Labor Economics*, 24:729–760.

Parsons, D. O. (1991). The health and earnings of rejected disability insurance applicants: Comment. *The American Economic Review*, 81(5):1419–1426.

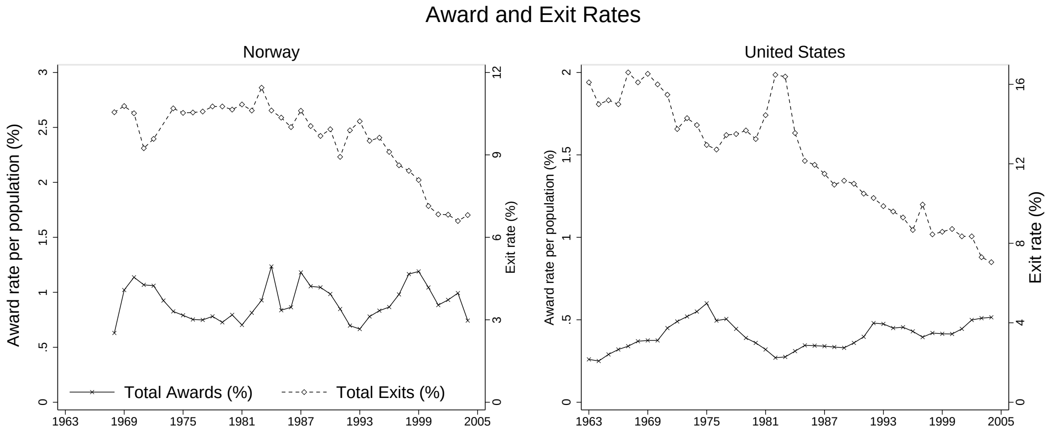
Pepper, J. (2000). The intergenerational transmission of welfare receipt: A nonparametric bounds analysis. *The Review of Economics and Statistics*, 82:472–288.

Rege, M., Telle, K., and Votruba, M. (2009). The effect of plant downsizing on disability pension utilization. *Journal of the European Economic Association*, 7(4):754–785.

Von Wachter, T., Song, J., and Manchester, J. (2011). Trends in employment and earnings of allowed and rejected applicants to the social security disability insurance program. *American Economic Review*, Forthcoming.

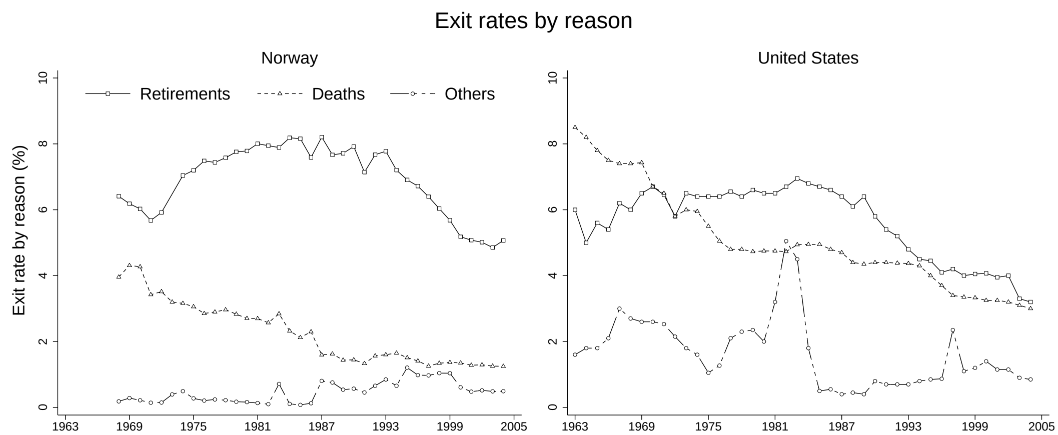
Appendix

Figure A.1:



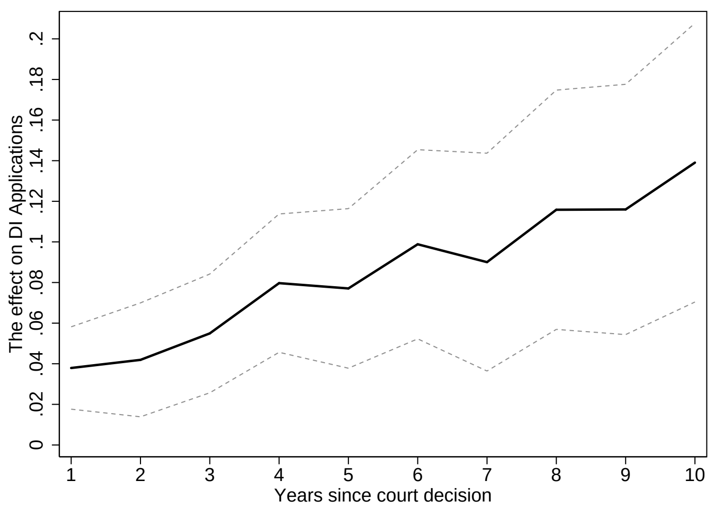
Note: The U.S. trends are based on Autor and Duggan (2006), while the Norwegian trends are collected from various issues of the SSA Supplement. The graphs show award rates in the insured population and exit rates from the DI program in both countries.

Figure A.2:



Note: 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.3: Estimates of DI Applications over Time.



Notes: Solid line is the yearly estimated effect of a parent's DI participation, instrumented with judge leniency, on their child's application for DI benefits. Data restricted to pre-2000 appeals so as to have a balanced 10 year sample. Dashed lines represent 95 percent confidence intervals.

Table A.1: Descriptive Statistics

Characteristic:	DI applicants Mean	DI applicants St. Dev.	DI appellants Mean	DI appellants St. 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]	0.0727	[0.2597]
Children below age 18	0.422	[0.494]	0.427	[0.495]
Previous earnings ($), 1-10 years prior to decision	30,559	[22,263]	18,458	[19,179]
Years of work, 1-10 years prior to decision	8.0	[3.1]	6.0	[4.0]
Mental disorders	N/A	N/A	0.305	[0.460]
Musculoskeletal disorders	N/A	N/A	0.397	[0.489]
DI allowed	0.715	[0.451]	0.113	[0.317]
Number of parents	97,623		7,414	
Panel B: Children				
Age (time of decision)	25.4	[4.6]	25.0	[4.6]
Female	0.463	[0.499]	0.487	[0.500]
Married	0.156	[0.363]	0.164	[0.371]
Foreign born	0.097	[0.296]	0.127	[0.333]
High school degree	0.403	[0.491]	0.367	[0.482]
College attendance	0.172	[0.377]	0.117	[0.321]
Children below age 18	0.358	[0.479]	0.31	[0.463]
Previous earnings ($), 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		14,894	

Notes: This table shows descriptive statistics for parents and their children for all applicants during the period 1992-2005 and all 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 beofre appeal in the DI appellant sample and the year before decision in the DI applicant sample. Unless stated otherwise, all parent and child characteristics are measured the year before the parent apply/appeal.

Table A.2: Estimates of child DI applications

	First stage		Reduced form		2SLS		N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: *Child applied for DI five years after parent's appeal*							
Parent allowed DI	0.918***	0.873***	0.057***	0.053**	0.062***	0.061***	14,894
	(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*							
Parent allowed DI	0.918***	0.873***	0.117***	0.112***	0.128***	0.128***	14,894
	(0.113)	(0.115)	(0.035)	(0.032)	(0.037)	(0.036)	
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.

Notes: Standard errors clustered at the judge level.

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

Characteristics	Norway	U.S.
Difficult to verify disorder	59.17 %	57.29 %
Other	40.83 %	42.71 %
Age (at decision on initial application)	52.2	49.11
Prior earnings relative to the median	70.98 %	69.91 %

Notes: This table illustrates differences between new recipients in Norway and in the U.S. The U.S. numbers are from Maestas et. al (2012), and the Norwegian numbers are drawn from the sample of awarded DI applicants during the years 2000-2003. Difficult to verify equals musculosceletal and mental disorder.Prior earnings are measured as AIE 3-5 years before the year of application/appeal.

Table A.4: Characteristics of DI Applicants and Appellants in Norway and the U.S.

	Norway		U.S.	
Characteristics	Applicants	Appellants	Applicants	Appellants
Difficult to verify disorder	60.88 %	69.69 %	58.49 %	62.18 %
Age (at decision on initial application)	51.1	47.1	47.1	46.1
Prior earnings relative to the median	66.54 %	50.41 %	60.45 %	56.25 %

Notes: This table illustrates differences between all applicants and all appellants in Norway and in the U.S. The U.S. numbers are from Maestas et. al (2012), and the Norwegian numbers are drawn from the sample of applicants during the years 2000-2003. Difficult to verify disorder includes musculosceletal and mental diagnoses. Prior earnings are measured as AIE 3-5 years before the year of application/appeal.

Table A.5: Characteristics of Marginal Applicants.

Characteristics:	First stage	$Pr[X = x]$	$Pr[X = x \mid complier]$	$\frac{Pr[X=x \mid complier]}{Pr[X=x]}$
Low education	1.035***	0.555	0.658	1.19
	(0.137)			
High education	0.706***	0.445	0.360	0.81
	(0.135)			
Young	0.876***	0.541	0.543	1.00
	(0.137)			
Old	0.880***	0.459	0.463	1.01
	(0.158)			
Married	0.922***	0.684	0.723	1.06
	(0.121)			
Not married	0.778***	0.316	0.281	0.89
	(0.166)			
High labor market experience	0.977***	0.434	0.486	1.12
	(0.184)			
Low labor market experience	0.813***	0.566	0.528	0.93
	(0.112)			

***p<.01, **p<.05, *p<.10.

Notes: Standard errors are clustered at the judge level.All characteristics are defined one year before the appeal.