How Disability Benefits in Early Life Affect Long-Term Outcomes*

Manasi Deshpande

August 27, 2020

Draft - Please do not circulate

Abstract

The debate over the Supplemental Security Income program for children reflects a key tradeoff in welfare programs: transfers to disadvantaged households could promote children's human capital development by increasing household resources, but conditioning those transfers on child health and family income could potentially discourage human capital development by creating perverse incentives. In this paper, I use two regression discontinuity designs (RDD) paired with Social Security administrative data to estimate the net effect of receiving SSI in childhood on adult earnings and to separately identify the household resources channel and perverse incentives channels. Using the first RDD, I find that removing children from SSI has a statistically insignificant net effect on child earnings in adulthood. Using the second RDD and a novel data linkage procedure to identify younger siblings in SSA administrative data, I find that removing youth from SSI at the age of 18 reduces the adult earnings of their younger siblings by about $5,000 annually. This finding suggests that SSI's household resources channel has a large positive effect on children's human capital development. I develop a decomposition procedure to determine the relative contributions of the income transfer and the perverse incentives channels to the net effect of SSI.

*University of Chicago and NBER, mdeshpande@uchicago.edu. I thank Linda Martin, Samuel Foster, William Lancaster, and Obie Blackmon of the Office of Data Development (Office of Research, Evaluation, and Statistics) at the Social Security Administration for providing data for this project. David Price provided helpful suggestions on the sibling linkage procedure. I thank Alessandra Voena and Jiada Ye for helpful comments. I thank Jason Weitze and Jiada Ye for excellent research assistance. I am grateful to the Ronzetti Initiative for the Study of Labor Markets at the Becker-Friedman Institute for financial support. This research was supported by the U.S. Social Security Administration through grant RDR18000003 from the US Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium. The findings and conclusions expressed are solely those of the author(s) and do not represent the views of SSA, any agency of the Federal Government, or the NBER.

1 Introduction

The debate over the Supplemental Security Income (SSI) program for children reflects some of the most contentious issues in the design of welfare programs. SSI provides cash benefits and Medicaid to 1.3 million low-income children with physical, mental, and behavioral disabilities in the United States. Supporters of the SSI program argue that SSI can improve the long-term outcomes of children by increasing the household resources of poor families raising a child with a disability. Depending on how the family spends the transfer, SSI could either increase household income or increase the amount of time a parent can spend with the child. Supporters call for expanding the SSI children's program and making it more generous.[1] Critics of the SSI program argue that conditioning benefits on the child's disability creates perverse incentives for families to present their children as disabled, which could discourage the child's educational achievement and human capital development. These critics argue for defunding the program or replacing the cash benefit with in-kind services.[2]

The optimal design of this program depends crucially on the relative magnitudes of these competing channels: improvements in child human capital through a "household resources" channel versus discouragement of child human capital through a "perverse incentives" channel. For example, if the household resources channel is large (small) relative to the perverse incentives channel, then making SSI more generous would likely improve (worsen) children's outcomes. Yet there is little evidence on the net effect of receiving SSI in childhood on long-term outcomes, let alone on the magnitudes of the channels.

In this paper, I estimate the net effect of receiving SSI in childhood on adult earnings and then decompose the net effect into the household resources channel and the perverse incentives channel. The household resources channel is expected to increase a child's adult earnings by increasing household income or parental time with the child and thereby promoting human capital development. The perverse incentives channel is expected to decrease a child's adult earnings by discouraging educational achievement and human capital development. Therefore the theoretical net effect of receiving SSI in childhood on adult earnings is ambiguous.

The ideal experiment to distinguish between the household resources channel and the perverse incentives channel would be to randomly assign families with children with disabilities to receive no benefit (Control), the standard SSI benefit (Standard Treatment), or an SSI

[1] See, e.g., Rebecca Vallas and Shawn Fremstad, "Maintaining and Strengthening Supplemental Security Income for Children with Disabilities," Center for American Progress, September 10, 2012; and Clare Proctor, "Advocates seek increase to Supplemental Security Income to benefit people with disabilities," *Chicago Sun-Times*, August 4, 2020).

[2] See, e.g., Patricia Wen, "The Other Welfare," *Boston Globe*, December 12, 2010; and Nicholas Kristof, "Profiting From a Child's Illiteracy," *New York Times*, December 7, 2012.

benefit that is not conditioned on the child's disability status (Unconditional Treatment). The difference in adult earnings (and other outcomes) between children in the Standard Treatment group and the Control group would provide an estimate of the net effect of SSI. The difference in adult outcomes between children in the Unconditional Treatment group and the Control group would provide an estimate of the household resources effect. The difference in adult outcomes between children in the Standard Treatment group and Unconditional Treatment group would provide an estimate of the perverse incentives effect. In lieu of this ideal experiment, I use two regression discontinuity (RD) designs paired with Social Security Administration (SSA) administrative data that together, under certain assumptions, identify each of these parameters.

To estimate the net effect of SSI on the child's adult earnings, I use a regression discontinuity design in SSI award date based on a change in SSA's policy on child medical reviews in 2004 ("Child RD").[3] Nearly all children who were awarded SSI just before FY2001 received a medical review three years later, while only a fraction of children who were awarded SSI just after FY2001 received a medical review due to an administrative budget cut. This discontinuity in the likelihood of receiving a medical review created a discontinuity in the likelihood of continuing to receive SSI benefits during childhood, which allows me to estimate the effect of (losing) SSI on adult earnings.

Using the Child RD, I find that removing children from SSI during childhood does not have a statistically significant effect on their earnings in adulthood. I estimate that children who are more likely to be removed in childhood have $185 higher annual adult earnings ($3,400 when scaled up by the likelihood of removal), but this estimate is not statistically different from zero. It also reflects in part a lower likelihood of receiving SSI in adulthood, which is expected to increase earnings through a contemporaneous income effect rather than a human capital effect. I expect this estimate to become more precise as more of the children enter prime working years and additional years of earnings data become available.

To identify the household resources channel separately from the perverse incentives channel, I use a regression discontinuity design in child birthdate based on a change in SSA's policy on age 18 medical reviews in 1996 ("Age-18 RD").[4] As a result of the Personal Responsibility and Work Opportunity Act of 1996, SSA was required to redetermine the eligibility of all children when they turned 18. Very few of the children who had an 18th birthday just before the date of PRWORA enactment—August 26, 1996—received an age 18 medical review, while nearly all children who an 18th birthday just after that date received a review. This

[3]The Child RD was originally used in Deshpande (2016b) to estimate the effect of removing children from SSI on their parents' earnings.

[4]The Age-18 RD was originally used in Deshpande (2016a) to estimate the effects of removing 18-year-olds from SSI on their own earnings in adulthood.

discontinuity in the likelihood of receiving an age 18 medical review created a discontinuity in the likelihood of receiving SSI benefits after age 18. I use the Age-18 RD to estimate the effect of removing an 18-year-old from SSI on the adult earnings of the *younger siblings* of the 18-year-old. This strategy arguably isolates the household resources channel: younger siblings are directly affected by the reduction in household resources, but they are not directly affected by the reversal of perverse incentives. In particular, because the benefit is tied to the 18-year-old's health and human capital, and not to the younger sibling's health and human capital, there is no change in the incentives facing the younger sibling when the 18-year-old is removed. I use two separate sibling samples: one is a sample of younger siblings known in SSA administrative because they also receive SSI, and the other is a sample of likely younger siblings that do not receive SSI. For the latter sample, I develop a novel probabilistic matching procedure in SSA administrative data based on parent names as well as child dates and places of birth. This is the first paper of which I am aware to develop a matching procedure to identify siblings in SSA administrative data.

Using the Age-18 RD, I find that removing 18-year-olds from SSI reduces the adult earnings of their younger siblings considerably. In both samples of younger siblings—those known in administrative data and receiving SSI, and those identified by the probabilistic matching procedure—the 18-year-old's removal reduces the younger siblings' adult earnings by \$5,000–\$6,000 annually between the ages of 19 and 30. This estimate suggests that SSI's household resources channel has a large positive effect on child human capital.

By estimating both the net effect of SSI and the "household resources" effect, I can separately identify the household resources effect relative to the perverse incentives effect. I develop a decomposition procedure to separate the net effect of SSI on child human capital into separate channels, using the empirical estimates from this paper as inputs. When the estimate of the net effect from the Child RD becomes more precise, this decomposition will allow me to formally decompose the net effect into the household resources and perverse incentives channels.

I can also further decompose the household resources effect into the part from household income and the part from parental time by estimating the effect of receiving SSI on parental earnings from the Age-18 RD. I find that parents do not change their earnings in response to the SSI loss, which suggests that in this case the household resources channel operates through household income rather than parent time spent on childcare.

This paper provides some of the first evidence on the effect of SSI on long-term child outcomes and the mechanisms through which this effect operates. Deshpande (2016a) estimates the effect of removing young adults from SSI at the age of 18 and finds that the removed youth have somewhat higher earnings but lower total income than those that are not removed.

But since the variation in removal does not occur until the age of 18, that paper does not address how SSI affects child development during the formative years of childhood, which is the most contentious issue in the debate over childhood SSI. Coe and Rutledge (2013) and Levere (2019) use variation from the 1990 Supreme Court decision that allowed children with mental and behavioral conditions qualify. Coe and Rutledge (2013) find that SSI improves child outcomes in adulthood, while Levere (2019) find that SSI worsens child outcomes in adulthood. There is no evidence of which I am aware on the separate channels through which SSI in childhood might affect adult outcomes, although there is evidence on the effects of unconditional cash transfers on child outcomes (Aizer et al. (2016)).

The paper proceeds as follows. Section 2 presents a conceptual framework for the effect of SSI in childhood on adult outcomes. Section 3 describes the empirical strategies and data. Section 4 presents the empirical results and the decomposition of the net effect into the household resources and perverse incentives channels. Section 5 concludes.

2 Conceptual Framework

The goal of this paper is to estimate the net effect of receiving SSI in childhood on the long-term outcomes of children and to decompose the net effect into different channels. The net effect here is the SSI program treatment effect: how SSI affects child human capital development and long-term outcomes. Estimating the net effect involves the classic challenge of separating the program treatment effect from selection. Since SSI is a means-tested program for children with disabilities, children who receive SSI are likely by design to have lower average potential earnings than children who do not receive SSI. The selection effect is expected to create a negative correlation between SSI receipt and adult earnings, which could lead to either an overestimate or underestimate of the treatment effect, depending on the sign of the treatment effect.

The net effect itself comprises multiple, potentially offsetting, channels through which SSI affects long-term outcomes (expected effect on adult earnings in parentheses):

Channel 1: Household resources (+). SSI provides cash benefits of around $700 per month to the families of children with qualifying conditions that have low income and assets. Given that the average household income of a children receiving SSI benefits is $35,000, the SSI benefit represents a substantial fraction of household income.[5] SSI also provides Medicaid

[5]Average household income among families with a child receiving SSI calculated from the Survey of Income and Program Participation.

to child and adult recipients. If parents do not change their labor supply behavior, then SSI will increase household income (channel 1a). If parents reduce their labor supply in response to the SSI benefits and use that additional time on child care, then SSI will increase parental time with the child (channel 1b). Of course, parents could also respond by reducing other sources of income, like support from family and friends. This still represents an expansion in the household budget set, though it is less clear whether we expect such a response to have direct effects on the child's outcomes. The part of SSI operating through the household resources channel is expected to increase a child's adult earnings.

Channel 2: Perverse incentives to qualify (-). Children qualify for SSI if they have a severe disability and if their parents have low income and assets. The definition of severity has changed over time, most notably in 1990 when a Supreme Court decision allowed children to qualify with mental and behavioral conditions like ADHD and autism. SSA conducts periodic medical reviews to verify that recipients still have a severe disability. If SSA demonstrates that the child has medically improved since the last review, then SSA can terminate the child's benefits. SSA also conducts periodic reviews of non-medical criteria like household income.[6] Critics argue that by conditioning welfare payments on the child's disability and family income, SSI might create perverse incentives for families to discourage achievement in their children. The part of SSI operating through the perverse incentives channel is expected to decrease a child's adult earnings.[7]

Channel 3: Adult SSI receipt (-). SSI receipt in childhood could increase the likelihood of SSI receipt in adulthood, which could decrease adult earnings through income and substitution effects in adulthood. SSI provides up to $700 per month in benefits, which could reduce adult earnings through an income effect. For adults, the SSI benefit is reduced by 50 cents for every $1 in earnings, which could reduce adult earnings through a substitution effect. Indeed, Deshpande (2016a) finds that removing 18-year-olds from SSI increases their adult earnings by about $3,000 annually, though not enough to replace the lost SSI income. If providing children with SSI benefits makes them more likely to receive SSI as adults, then childhood SSI receipt is expected to decrease a child's adult earnings. Unlike the household resources and perverse incentives channels, however, this effect operates not through child human capital, but through incentives in adulthood. For the purposes of this paper, this channel is therefore a "nuisance" parameter—one that must be estimated to identify

[6] The exception is medical reviews of children receiving SSI at age 18, which is treated as a de novo review.

[7] Another possible channel that is negatively correlated with long-term outcomes is expectations about future SSI benefits. If households incorrectly expect that children will continue receiving SSI benefits in adulthood, then SSI could discourage human capital investment in childhood, which would decrease adult earnings. I return to this channel in Section 4.

the channels of interest (household resources and perverse incentives), but is not directly relevant to the question of how SSI affects child human capital development.

The net effect can be decomposed into these three channels:

$$\begin{aligned}\text{Net Effect of SSI} = &\underbrace{\text{HH income}}_{\text{Channel 1a (+)}} + \underbrace{\text{Parent time}}_{\text{Channel 1b (+)}} \\ &+ \underbrace{\text{Perverse incentives}}_{\text{Channel 2 (-)}} + \underbrace{\text{Adult SSI}}_{\text{Channel 3 (-)}}\end{aligned} \tag{1}$$

Understanding the relative roles of these channels is critical to the optimal design of disability programs. Distinguishing between selection and program treatment effects is a first-order question in any policy evaluation. Within the treatment effect, quantifying the three channels separately—increased household resources (channel 1), perverse incentives to qualify (channel 2), and adult SSI receipt (channel 3)—can help to design the SSI program to minimize harms and maximize benefits. For example, if higher household resources improve child outcomes but perverse incentives have a minimal effect, then this suggests that expanding the program could improve child welfare substantially.

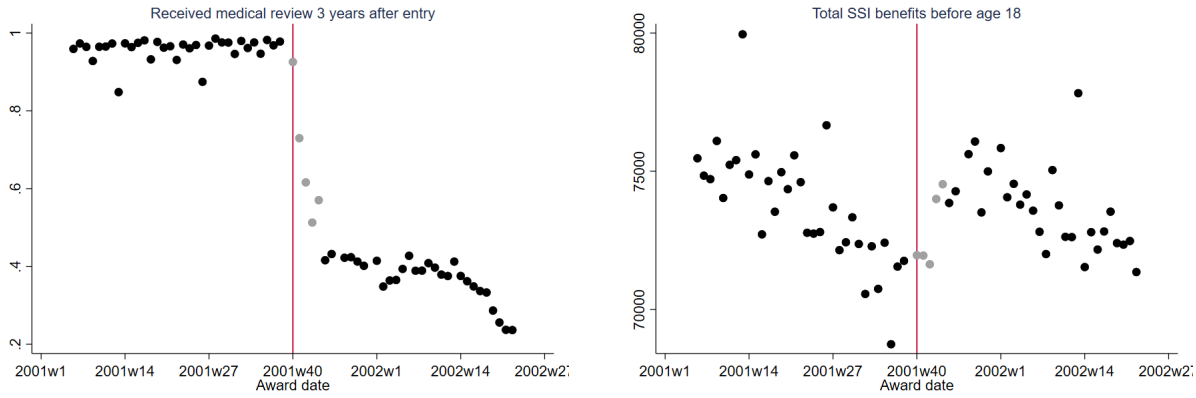
3 Empirical Strategies and Data

I use two primary empirical strategies paired with SSA administrative data to estimate the parameters in equation (1). The first empirical strategy ("Child RD," originally used in Deshpande (2016b)) is a regression discontinuity design in the date of a child's entry onto SSI. Due to an administrative budget cut in FY2005, children who entered SSI just after October 1, 2001, were much less likely to receive a medical review three years later than children who entered SSI just before this date (see Figure 1). This means that children who entered just before the cutoff were more likely to be removed from SSI, and they received lower SSI payments in childhood. Deshpande (2016b) uses this variation to study the effects of removing young children from SSI on their parents' earnings. In this paper, I use the Child RD strategy to estimate the effect of removing children from SSI on their adult earnings using the following specification:

$$Y_i = \alpha + \beta \text{FY2001}_i + \gamma \text{AwardDate}_i + \delta(\text{FY2001}_i \times \text{AwardDate}_i) + \varepsilon \tag{2}$$

where Y_i is child earnings in adulthood, FY2001_i is an indicator for having a date of birth before the FY2002 award date cutoff, and AwardDate_i is the award date running variable. The coefficient of interest is the β since it gives the effect on adult earnings of having an award date before the cutoff. The children in the original Child RD sample were between the ages of 3 and 12 in 2004, making them 17–26 years old in 2018, the latest available year of earnings data. Deshpande (2016b) presents covariate balance tests.

Figure 1: Child RD Empirical Design



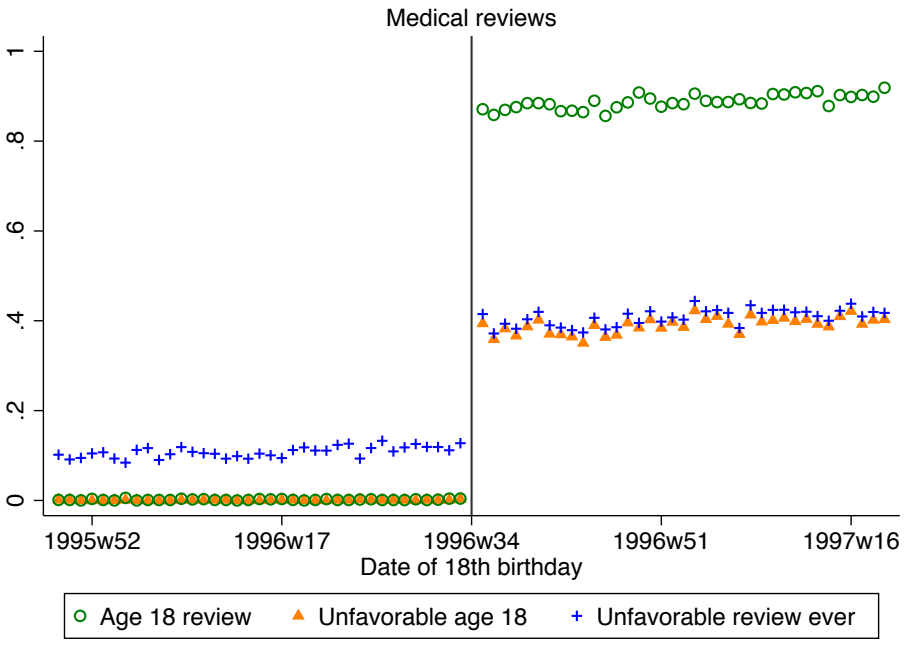
Notes: Figure shows the empirical design for the "Child RD" strategy. The left-hand-side graph plots the likelihood of receiving a medical review three years after entering SSI, by the date of SSI entry (award date). Due to an administrative budget cut, children who entered SSI just after October 1, 2001 (the start of FY2002) were much less likely to receive a medical review three years later than children who entered SSI prior to that date. The right-hand-side graph plots total SSI benefits received by date of SSI entry. The sample is SSI children with a three-year medical review schedule and an award date within 200 days of the FY2001/2002 cutoff. Data are binned by week of entry. The five gray dots to the right of the cutoff occur during the transition to lower medical review rates in FY2005.

The second empirical strategy ("Age-18 RD," originally used in Deshpande (2016a)) is a regression discontinuity design in child birthdate. Children receiving SSI benefits who turned 18 years old just after PRWORA enactment (August 22, 1996) were more likely to get a medical review and be removed from SSI at age 18 than those who turned 18 years old just before enactment (see Figure 2). Deshpande (2016a) uses this variation to study the effects of removing 18-year-olds from SSI on their adult earnings. Since the children in the original Age-18 RD sample were 18 years old in 1996, they turned 40 years old in 2018. In this paper, I use the Age-18 RD strategy to estimate the effect of removing children from SSI at the age of 18 on their younger siblings' earnings in adulthood. Throughout the paper, I use "18-year-olds" to refer to the original sample of children in the Age-18 RD and "younger siblings" to refer to the younger siblings of those 18-year-olds. I estimate the following specification:

$$Y_i = \alpha + \beta \text{AfterPRWORA}_i + \gamma \text{DOB}_i + \delta(\text{AfterPRWORA}_i \times \text{DOB}_i) + \varepsilon \tag{3}$$

where Y_i is younger sibling earnings in adulthood, AfterPRWORA$_i$ is an indicator for the 18-year-old having a date of birth after the date of PRWORA enactment (August 21, 1996), and DOB$_i$ is the 18-year-old date of birth running variable. The coefficient of interest, β, gives the effect on younger sibling earnings of the 18-year-old having a birthdate after the cutoff. To account for multiple siblings in the same household, I cluster standard errors at the 18-year-old level. Deshpande (2016a) presents covariate balance tests.

Figure 2: Age-18 RD Empirical Design



Notes: Figure plots proportion of SSI children in each birthweek bin who receive an age 18 medical review, receive an unfavorable age 18 medical review, and ever receive an unfavorable medical review (through 2013). Sample is SSI children with an 18th birthday within 37 weeks of the August 22, 1996, cutoff.

I use SSA administrative data for the Child RD, Age-18 RD, and all related analyses in this paper. The Supplemental Security Record (SSR) includes demographic information and detailed monthly benefit history information for every individual who has ever received SSI benefits. After constructing the Child RD and Age-18 RD samples from the SSR, I link the individuals in these samples to annual earnings from the Master Earnings File, and to future SSI and SSDI applications in the F831 files, and to future SSI and SSDI receipt through the SSR and the Master Beneficiary Record, respectively. For children, the SSR also includes identifiers for their parents, including Social Security Number, full name, sex, and date of birth. Using this link between children and parents, I can observe annual parent earnings through the Master Earnings File. In addition, I use this link between children and parents to identify siblings of the children, described in more detail below and in Appendix A.

I identify each parameter in equation (1) using either the Child RD or the Age-18 RD paired

with a specific outcome. Note that both of these RDs involve the effect of removing (or continuing) children who were previously on SSI, not the effect of awarding benefits to a child who had never previously received SSI. For the purposes of this paper, and especially the decomposition exercise, I assume that the effect of an additional year of SSI is the same regardless of whether the child was previously receiving SSI or not. Although this assumption simplifies the analysis considerably, it is possible that it does not hold. I identify each parameter using the following strategies:

Net effect: Child RD on adult earnings. To estimate the net effect of SSI in childhood on long-term outcomes, I use the Child RD with the original sample of children linked to their adult earnings. Recall that all of the children on the left-hand-side of the discontinuity in Figure 1 received a medical review on time, while the children on the right-hand-side of the discontinuity were less likely to receive a medical review on time and therefore more likely to remain on SSI during childhood. This strategy therefore allows me to estimate the net effect of receiving SSI in childhood on adult earnings. The children in the Child RD sample were 17–26 years old in 2018, the latest available year of earnings data, making a sufficient number old enough for their observed earnings to be meaningful representations of their labor market success.

Household resources (channel 1): Age-18 RD on younger sibling earnings. To isolate the effect of SSI on long-term outcomes operating through increased household resources, I use the Age-18 RD linked to the adult earnings of the *younger siblings* of the 18-year-olds in the original sample. The hypothesis motivating this analysis is that younger siblings (especially if they don't receive SSI themselves) are affected by the 18-year-old's removal from SSI through the household resources channel, but not through perverse incentives created by the program because the SSI income is not conditioned on the younger sibling's health or human capital. If this hypothesis is true, then studying effects on younger siblings allows me to isolate the household resources channel (channel 1) and distinguish it from the perverse incentives channel (channel 2). I use two different sibling samples in this analysis:

1. Sample 1: Younger siblings who themselves receive SSI: Identifying younger siblings who also receive SSI is straightforward since the Supplemental Security Record includes parent SSN. I define siblings in the SSR as children who have at least one parent in common on their record. Because of the ability to match precisely, this sample has the advantage of no or few false sibling matches. The disadvantage of this sample is that younger siblings who also receive SSI during childhood might be affected by the

18-year-old's review through expectations about the younger sibling's own likelihood of receiving benefits in adulthood. Moreover, the direction of the expectations effect is ambiguous: younger siblings could increase their expectations of receiving SSI in adulthood if the 18-year-old survives the review, or decrease their expectations if the 18-year-old is removed). For this reason, I use a second, broader sample of younger siblings.

2. Sample 2: Younger siblings who do not receive SSI: Because of interpretation issues created by studying siblings who themselves receive SSI in adulthood, I construct a sample of younger siblings who do not receive SSI benefits using a probabilistic matching procedure in SSA administrative. Unlike IRS data, SSA administrative data does not provide information on family relationships for individuals who do not receive benefits from SSA, making it much more difficult to identify siblings who do not receive SSI benefits. To the best of my knowledge, no one has successfully identified sibling linkages using SSA administrative data.[8] I develop a new matching procedure using the Numident that matches siblings on mother's full name and father's full name. I use other observables to determine which "potential" siblings are likely to be true matches: e.g., place of birth, dates of birth of the potential sibling relative to the 18-year-old, and date of birth of the younger sibling relative to the mother of the 18-year-old. Appendix A provides the details of this novel data linkage procedure. The sample includes "potential" siblings from this matching procedure who have a relatively high likelihood of being a correct match, who are at least four years younger than the 18-year-old (so that there is time to experience the effects of the 18-year-old's removal), and who do not themselves receive SSI benefits as children (i.e., do not appear in Sample 1). The advantage of Sample 2 relative to Sample 1 is that the perverse incentives and expectations channels are likely to be even more muted for siblings who do not receive SSI themselves in childhood. Therefore this sample is likely to isolate the household resources channel (channel 1) even more so than the first sample. The disadvantage of this sample is the high rate of false positives and false negatives: it will include some "potential siblings" from the probabilistic match who are not true siblings, and it will exclude a large number of true siblings. In addition, the sample disproportionately includes children from stable families since those siblings are more likely to match (since, e.g., two parent names match, or mother is more likely to have the same last name for different births). By construction, this sample is also less representative of children who themselves receive SSI benefits.

[8] Price and Song (2020)

Household income (channel 1a) and parent time (channel 1b): Age-18 RD on parent earnings. From equation (1), the household resources channel itself comprises two sub-channels: household income and parent time. Estimating the effect of the Age-18 RD on parent earnings can shed light on which of these sub-channels dominates. If parents increase their earnings in response to the loss of the 18-year-old's SSI income, then the effect of SSI on younger sibling earnings likely operates through parent time spent on childcare rather than through household income. If parental earnings do not respond, then the loss of the 18-year-old's SSI income will reduce household income, and the effects on younger siblings might operate through that reduction in household income.

Adult SSI receipt (channel 3): Child RD on adult SSI receipt. Although the channel of adult SSI receipt is not of direct interest in child human capital development, estimating this parameter is necessary in order to identify the parameters of interest in equation (1). To do this, I use the Child RD with the original sample of children linked to their SSI receipt in adulthood. I do not expect this effect to be large because all children receiving SSI benefits get a de novo adult medical review at age 18 that is stricter than the child medical review. Therefore, the children who entered just after the FY2002 cutoff and were spared from medical review as children would later receive the stricter age 18 medical review. The age 18 review process is likely to even out adult SSI enrollment after age 18 across the cutoff.

Perverse incentives (channel 2): Residual. Neither the Child RD nor the Age-18 RD identifies the perverse incentives channel on its own. I instead estimate this channel as the residual of the other channels using equation (1). I develop a decomposition procedure to estimate the perverse incentives channel in Section 4.5.

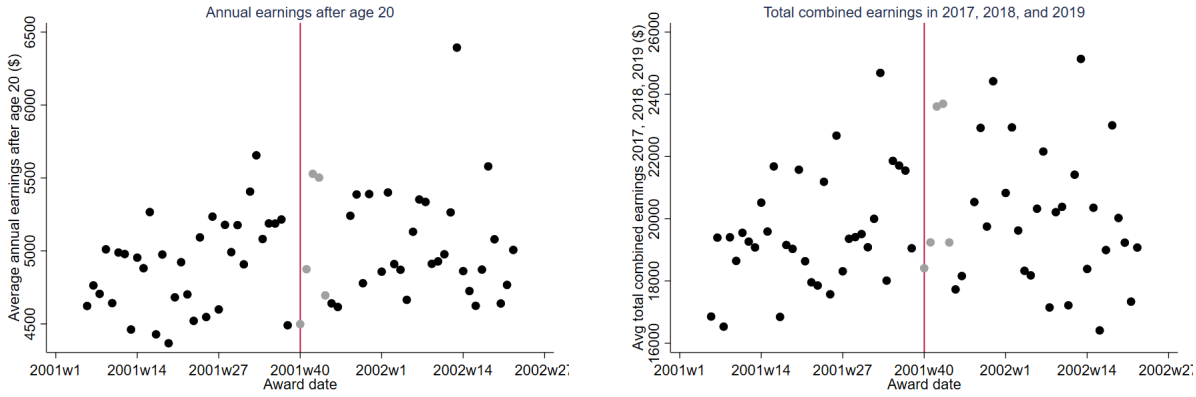
4 Empirical Results

4.1 Net effect: Child RD on adult earnings

Figure 1 shows the first-stage effect of entering SSI before the cutoff from the Child RD, replicated from Deshpande (2016b). Children who enter just before the cutoff are 50 percentage points more likely to get a medical review 3 years later (see Table 1). As a result, they are 5.8 percentage points less likely to be on SSI and receive $5,100 less in total SSI payments before age 18. When this figure is scaled up by the likelihood of being on SSI

before age 18, children who are removed as a result of the review lose on average $87,200 (= 5100/0.058) in SSI payments during childhood.

Figure 3: Effect of Child RD on Adult Earnings



Notes: Figure shows adult earnings for the "Child RD" strategy. The left-hand-side graph plots average annual earnings after age 20 by award date. The right-hand-side graph plots total combined earnings in 2017, 2018, and 2019 by award date. The sample is SSI children who have a three-year medical review schedule, have an award date within 200 days of the FY2001/2002 cutoff, and reach age 20 by 2018, the latest available year of earnings data. Data are binned by week of entry. The five gray dots to the right of the cutoff occur during the transition to lower medical review rates in FY2005.

The new finding in this paper is the effect of the Child RD on the adult earnings of the children in the original sample. Figure 3 shows the reduced form effect of entering SSI before the cutoff on adult earnings from the Child RD. Children who enter just before the cutoff and are more likely to be removed from SSI as children have $185 higher annual earnings after age 20 relative to children who enter just after the cutoff (see Table 1). However, the estimate is not statistically significant and Figure 3 does not show an obvious discontinuity. When scaled by the the likelihood of being on SSI before age 18, the reduced-form estimate suggests that removing children increases their annual adult earnings by $3,400, but again, this estimate is imprecise. I expect the estimate to become more precise with additional years of data, as more of the children enter prime working years.

4.2 Household resources (channel 1): Age-18 RD on younger sibling earnings

Figure 2 shows the first-stage effect of having an 18th birthday after the cutoff from the Age-18 RD, replicated from Deshpande (2016a). Children who have an 18th birthday just after the date of PRWORA enactment were 87 percentage points more likely to receive a medical review and 37 percentage points more likely to receive an unfavorable receive an unfavorable review than those whose birthday came just before the cutoff (see Table 2).

Table 1: Child RD: First stage, reduced form, and IV estimates

	First stage/reduced form		IV estimates		RHS
	Pt Est	Std Err	Pt Est	Std Err	Mean
First stage (N = 36,388)					
Received review	0.495***	(0.00984)			0.30
On SSI from 2004 to age 18	-0.0584***	(0.00662)			0.89
Avg ann SSI pay before 18	-349.9***	(44.33)	5,989***	(644.4)	$5,029
Total SSI pay before 18	-5,092***	(557.4)	-87,161***	(6,586)	$73,158
Child outcomes (N = 29,173)					
On SSI after age 20	-0.0112	(0.0130)	-0.206	(0.231)	0.47
Earnings after age 20	184.9	(211.2)	3,405	(3,845)	$5,021
Parent outcomes (N = 33,076)					
Mother's earnings after 2004	832.4***	(255.9)	14,812***	(4,501)	$7,519
Household earnings after 2004	717.4*	(419.8)	13,063*	(7,336)	$14,072

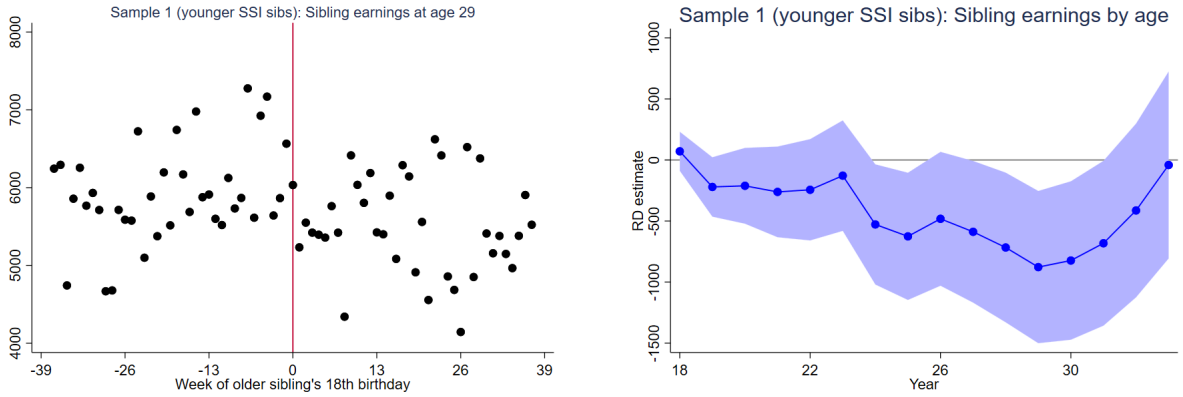
Notes: Table presents regression discontinuity estimates for the "Child RD" strategy. The sample is SSI children with a three-year medical review schedule and an award date within 200 days of the FY2001/2002 cutoff. Sample size for child outcomes is lower than the first stage sample size because not all children have reached age 20 by 2018, the latest available year of earnings data. Sample size for parent outcomes is low than first stage sample size because not all children have a parent. All IV estimates use "On SSI from 2004 to age 18" as the endogenous variable.

The new finding in this paper is the effect of the Age-18 RD on the adult earnings of younger siblings of the 18-year-olds in the original sample. Recall that I construct two samples of younger siblings: Sample 1 contains verified siblings who themselves receive SSI as children, while Sample 2 contains likely (probabilistically matched) siblings who do not receive SSI. Figure 4 shows the reduced form effect on adult earnings for Sample 1. The left-hand-side graph shows sibling earnings for Sample 1 at age 29: the younger siblings of the 18-year-olds with birthdates just after the cutoff earn $880 less at age 29 (significant at 1% level). When scaled up by the likelihood that the 18-year-old is on SSI for the remainder of the younger sibling's childhood (until age 18), this estimate suggests that removing an 18-year-old from SSI reduces these younger siblings' earnings by $7,600 at age 29. The average annual estimate from ages 19 to 30 is −$5,000 (significant at 5% level). The RHS graph shows the RD estimate for sibling earnings at each age from 18 to 33. The earnings gap appears to grow over time, probably because the level of earnings increases over the life cycle.[9] The estimates in Table 2 indicate that the effect on younger sibling earnings operates at both

[9]The attenuation in the last few years is the result of an imbalanced panel: only the oldest of the younger siblings remain in the sample at the older ages. The effects are smallest for the oldest of the younger siblings because they are least exposed to the 18-year-old's removal, so the effect attenuates over time. When I limit to the younger of the younger siblings, the estimates are less precise in all years but the effect does not attenuate.

the extensive and intensive margins: removing the 18-year-old reduces the likelihood that the younger sibling has any earnings after age 30 by 33 percentage points and the likelihood that the younger sibling earns above $20,000 by 18 percentage points.

Figure 4: Effect of Age-18 RD on Younger Sibling Adult Earnings (Sample 1)



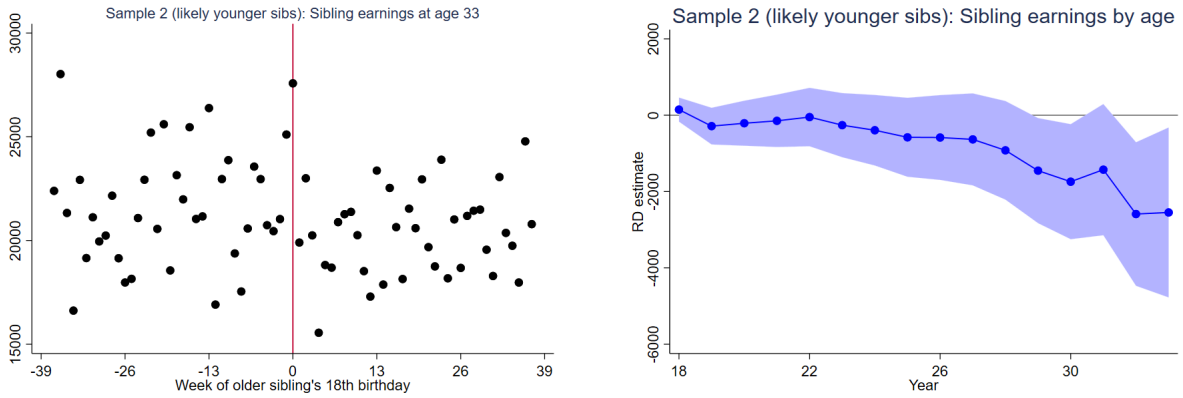
Notes: Figure shows adult earnings of known younger siblings of the 18-year-olds in the Age-18 RD sample. The left-hand-side graph plots the earnings of younger siblings in Sample 1 at age 29 by the week of the older sibling's 18th birthday. Sample 1 comprises children in the Supplemental Security Record who have at least one parent in common with an 18-year-old in the original Age-18 RD sample (SSI children with an 18th birthday within 37 weeks of the August 22, 1996, cutoff), are younger than the 18-year-old, and received SSI at some point in their childhood. The right-hand-side graph plots the regression discontinuity estimate at each age from 18 to 33. The attenuation in the last few years is the result of an imbalanced panel: only the oldest of the younger siblings remain in the sample at the older ages. The effects are smallest for the oldest of the younger siblings because they are least exposed to the 18-year-old's removal, so the effect attenuates over time.

Figure 5 shows the reduced form effect on adult earnings for Sample 2, the sample of a broader set of siblings who do not themselves receive SSI in childhood. These estimates are noisier, probably because the sample includes a nontrivial fraction of false sibling matches. However, even these younger siblings experience an effect of the 18-year-old's review, especially later in adulthood. The left-hand-side graph shows sibling earnings at age 33. Younger siblings of 18-year-olds with birthdates just after the cutoff earn $2,500 less at age 33 (significant at 5% level) (see Table 2). When scaled up by the likelihood that the 18-year-old is on SSI for the remainder of the younger sibling's childhood, this estimate suggests that removing an 18-year-old from SSI reduces these younger siblings' earnings by $14,600. The average annual estimate from 19 to 30 is −$5,800 and after age 30 is −$11,600. The right-hand-side graph shows the RD estimate for sibling earnings at each age from 18 to 33. The earnings gap appears to grow over time for this sample as well.[10] The effect on younger sibling earnings is larger for Sample 2 than for Sample 1, probably because the Sample 2 younger siblings do not themselves receive SSI for a disability and therefore have higher earnings capacity.

[10]This increase does not appear to be driven by the oldest of the younger siblings in an imbalanced panel, since the effect grows at the older ages even when I restrict to relatively young siblings.

In contrast to Sample 1, the effects for Sample 2 appear to operate largely on the intensive margin.

Figure 5: Effect of Age-18 RD on Younger Sibling Adult Earnings (Sample 2)



Notes: Figure shows adult earnings of likely younger siblings of the 18-year-olds in the Age-18 RD sample. The left-hand-side graph plots the earnings of younger siblings in Sample 2 at age 33 by the week of the older sibling's 18th birthday. Sample 2 comprises children in the Numident who are likely to be a sibling of an 18-year-old in the original Age-18 RD sample (SSI children with an 18th birthday within 37 weeks of the August 22, 1996, cutoff), based on mother name, father name, place of birth, and dates of birth; who are at least four years younger than the 18-year-old; and who never received SSI in their childhood. See Appendix Section A for more details on the sibling matching procedure. The right-hand-side graph plots the regression discontinuity estimate at each age from 18 to 33.

Besides household resources, are there other possible explanations for the effect of the 18-year-old's removal on the adult earnings of younger siblings? First, as I discuss above, for the younger siblings who themselves receive SSI (Sample 1), it is possible that the 18-year-old's medical review affects the younger siblings' expectations about the likelihood that they themselves will be removed from SSI at age 18. This potential confound is the motivation for constructing the broader set of siblings in Sample 2 who do not receive SSI benefits themselves. Since I find large effects on adult earnings even for the Sample 2 siblings, it is unlikely that expectations about one's own SSI receipt in adulthood are driving the effect on younger sibling earnings. Second, I investigate whether the effect on younger sibling adult earnings could be operating through the younger sibling's own SSI receipt in adulthood. The effect of the 18-year-old's removal on the adult SSI receipt of the younger siblings in both samples is a precise zero (see Table 2). Third, it is possible that the 18-year-old's removal increases the likelihood that the younger sibling will have to support the 18-year-old when both are adults, and this extra demand on the younger sibling affects his or her earnings. While this is possible, I would expect that channel to have the effect of increasing younger sibling earnings, whereas I find here that the 18-year-old's removal decreases younger sibling earnings. Moreover, because many of the 18-year-olds in the original sample were removed from SSI some years later as a result of an adult medical review, there are relatively small

Table 2: Age-18 RD: First stage, reduced form, and IV estimates

	First stage/reduced form		IV estimates		LHS
	Pt Est	Std Err	Pt Est	Std Err	Mean
18-year-old outcomes (N = 81,800)					
Received age-18 review	0.866***	(0.00317)			0.00
Unfavorable age-18 review	0.372***	(0.00467)			0.00
On SSI after 18	-0.0815***	(0.00544)			0.52
Avg ann earn after 18	366.3***	(126.3)	4,495***	(1,442)	$5,017
Sample 1 sib outcomes (N = 22,332)					
18yo on SSI when younger sib < 18	-0.120***	(0.0112)			0.82
18yo SSI income when younger sib < 18	-1,170***	(107.2)	-9,763***	(418.9)	$6,294
On SSI after 1996	0.00658	(0.0104)	-0.0601	(0.0874)	0.50
Avg ann earn 19–30	-577.5**	(233.4)	-5,017**	(2,056)	$5,002
Avg ann earn after 30	-501.9	(358.7)	-4,314	(3,087)	$7,269
Avg ann earn after 30 > $0	-0.0386***	(0.0138)	-0.328***	(0.122)	0.42
Avg ann earn after 30 > $20K	-0.0209**	(0.00865)	-0.179**	(0.0764)	0.11
Avg ann earn after 30 > $30K	-0.00448	(0.00609)	-0.0387	(0.0520)	0.05
Sample 2 sib outcomes (N = 15,271)					
18yo on SSI when younger sib < 18	-0.149***	(0.0156)			0.73
18yo SSI income when younger sib < 18	-1,160***	(147.0)	-7,770***	(448.7)	$5,425
On SSI after 1996	0.000409	(0.00422)	-0.00274	(0.0283)	0.02
Avg ann earn 19–30	-876.3*	(490.9)	-5,815*	(3,271)	$14,106
Avg ann earn after 30	-1,744**	(787.9)	-11,573**	(5,287)	$20,680
Avg ann earn after 30 > $0	-0.0217	(0.0150)	-0.144	(0.101)	0.74
Avg ann earn after 30 > $20K	-0.0424**	(0.0173)	-0.281**	(0.117)	0.40
Avg ann earn after 30 > $30K	-0.0203	(0.0152)	-0.135	(0.101)	0.25
Mother earnings					
Full sample, 1997–2002	-18.24	(168.8)	-110.8	(1,025)	$8,652
Sample 1, when younger sib < 18	-352.0	(282.8)	-2,882	(2,349)	$5,106
Sample 2, when younger sib < 18	-354.6	(524.7)	-2,364	(3,519)	$9,673
Average annual earnings 21–25					
18-year-old	486.4***	(99.24)	2,724***	(526.1)	$3,860
Sample 1 sib	-386.1*	(203.8)	-3,291*	(1,730)	$4,301
Sample 2 sib	-285.0	(383.3)	-1,909	(2,561)	$11,159

Notes: Table presents regression discontinuity estimates for the "Age-18 RD" strategy. The sample for "18-year-old outcomes" is SSI children with an 18th birthday within 37 weeks of the August 22, 1996, cutoff. IV estimates for this sample use "On SSI after 18" as the endogenous variable for "Avg ann earn after 18," and use "On SSI 21-25" as the endogenous variable for "Avg ann earn 21–25." The sample for "Sample 1 sib outcomes" comprises children in the Supplemental Security Record who have at least one parent in common with an 18-year-old in the original Age-18 RD sample, are younger than the 18-year-old, and received SSI at some point in their childhood. The sample for "Sample 2 sib outcomes" comprises children in the Numident who are likely to be a sibling of an 18-year-old in the original Age-18 RD sample, based on mother name, father name, place of birth, and dates of birth; who are at least four years younger than the 18-year-old; and who never received SSI in their childhood. IV estimates for Sample 1 and Sample 2 siblings use "18yo on SSI when younger sib < 18" as the endogenous variable. Standard errors are clustered at the 18-year-old (household) level.

differences in the adult SSI receipt of the 18-year-olds: Appendix Figure 7 shows that the discontinuity for SSI receipt shrinks from 23 percentage points in 2000 to just 2.7 percentage

points in 2019. Given the attenuation of the first stage over time, the differences in demands on younger siblings on either side of the discontinuity are likely small.

4.3 Household income (channel 1a) and parent time (channel 1b): Age-18 RD on parent earnings

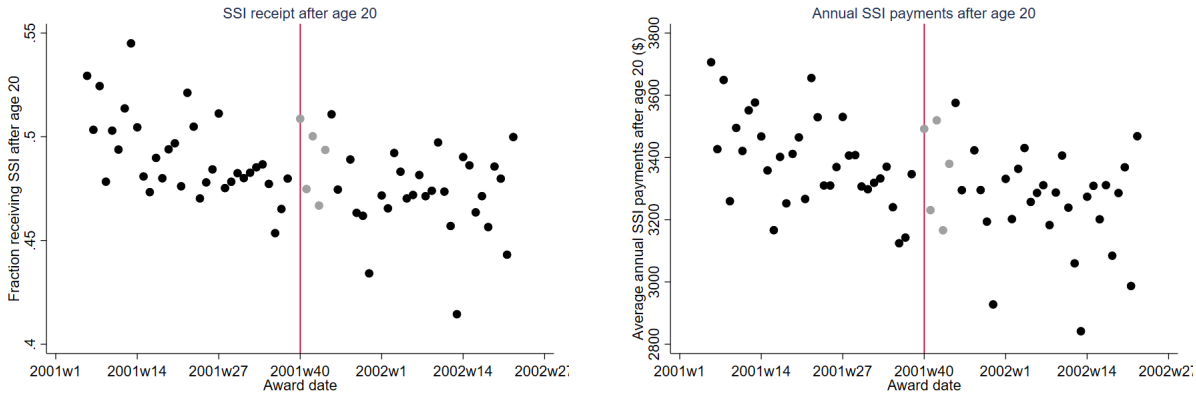
The effect of the Age-18 RD on parent earnings sheds light on the extent to which a child's removal from SSI affects household resources through reduced household income versus reduced parent time with the child. I focus on mother earnings since a large fraction of households are headed by single mothers, and because in two-parent households mother's labor supply is generally more responsive than father's labor supply. Table 2 presents estimates of the effect of the Age-18 RD on mother earnings for the original sample of 18-year-olds and for the Sample 1 and Sample 2 siblings. Mothers do not increase their annual earnings in response to the SSI loss in any of these groups. This finding suggests that the household experiences a net decline in income as a result of the 18-year-old's removal. If so, the adverse effect on younger sibling earnings likely operates through a decrease in household income rather than a decrease in parent time spent on child care.[11]

4.4 Adult SSI receipt (channel 3): Child RD on adult SSI receipt

To estimate the adult SSI receipt channel, I study the effect of the Child RD on adult SSI receipt. Recall that all children receive an age-18 medical review in the year or two after turning 18, and that this review has stricter standards than childhood reviews. Figure 6 presents adult SSI receipt for the original sample of children using an indicator of SSI receipt after age 20 (left-hand-side) and the annual amount of SSI payments after age 20 (right-hand-side). Children who enter just before the cutoff are 1.1 percentage points less likely to be on SSI and receive $130 less annually in SSI benefits after age 20 (not significant) (see Table 1). When scaled by the likelihood of receiving SSI before age 18, the reduced-form estimate suggests that removing children from SSI decreases the likelihood of SSI receipt in adulthood by 21 percentage points. Even though the estimate for adult SSI receipt is imprecise, this effect on adult SSI could partially account for the positive (imprecise) net effect of the Child RD on adult earnings. Therefore I need to account for this "nuisance parameter" in estimating the other parameters in equation (1).

[11] In contrast, Deshpande (2016b) finds that parents in the Child RD sample more than fully replace the lost SSI income with earnings, as shown in Table 1. The difference in the parental earnings responses in the Child RD and the Age-18 RD could be explained by parent age or child age. The children in the Child RD sample are younger (between 3 and 12 years at the time of removal) than the children in the Age-18 RD

Figure 6: Effect of Child RD on Adult SSI Receipt



Notes: Figure shows adult SSI receipt for the "Child RD" strategy. The left-hand-side graph plots the likelihood of receiving SSI benefits after age 20 by award date, since the vast majority of all age-18 medical reviews have been completed by then. The right-hand-side graph plots average annual SSI payments after age 20 by award date. The sample is SSI children who have a three-year medical review schedule, have an award date within 200 days of the FY2001/2002 cutoff, and reach age 20 by 2018, the latest available year of earnings data. Data are binned by week of entry. The five gray dots to the right of the cutoff occur during the transition to lower medical review rates in FY2005.

4.5 Decomposing the net effect

The empirical strategies and outcomes presented in the section above shed light on the net effect of SSI in childhood on adult earnings and the channels through which this effect operates. A formal decomposition analysis based on equation (1) can help estimate more precisely the relative contributions of each channel. In particular, the formal decomposition helps to identify the magnitude of the perverse incentives channel, since neither empirical strategy on its own identifies this channel. I start by using the estimates of the net effect, household resources, and adult SSI receipt to determine the contribution of the perverse incentives channel:

$$\text{Net Effect of SSI} = \underbrace{\text{HH resources}}_{\text{Channel 1}} + \underbrace{\text{Perverse incentives}}_{\text{Channel 2}} + \underbrace{\text{Adult SSI}}_{\text{Channel 3}} \tag{4}$$

Unfortunately, the empirical estimate of the net effect of receiving SSI in childhood on adult earnings from the Child RD is imprecise, and so the decomposition using the current estimates will not be meaningful. I expect the estimate of the net effect to become more precise with additional years of data. For now, for the purposes of exposition, I present the decomposition analysis taking the current point estimates at face value. However, it is important to note that the resulting estimates will be highly imprecise and to interpret them

sample (18 years old at the time of removal), and their parents are on average younger as well.

with caution. The point estimate for the net effect of removing a child from SSI on annual earnings is in their early 20s is (an imprecise) $3,400. For the household resources effect, I use the earnings of the Sample 1 siblings since they receive SSI themselves in childhood and are therefore more comparable to the Child RD sample than Sample 2 siblings. The point estimate for the annual earnings of the Sample 1 younger siblings from ages 21–25 is −$3,300 (see the last panel of Table 2). For the adult SSI receipt channel, I must multiply the effect of removal on SSI receipt by the contemporaneous effect of adult SSI receipt on adult earnings. The effect of removing a child from SSI on adult SSI receipt is −21 percentage points (from Table 1). Using the estimate of the effect of adult SSI receipt on contemporaneous adult earnings from the last panel of Table 2, the estimate of how SSI's adult SSI receipt channel affects earnings is $560 (= 0.21 × 2700). Therefore the estimate of how SSI's perverse incentives channel affects earnings is (a highly imprecise) −$6,200. Again, it is important to caution that this estimate is for exposition purposes only, and additional years of data will allow for a more precise estimate.

I further attempt to decompose the household resources channel into the sub-channels of household income (channel 1a) and parent time spent on child care (channel 1b) using additional assumptions:

$$\begin{aligned} \text{HH resources} &= \underbrace{\text{HH income}}_{\text{Channel 1a}} + \underbrace{\text{Parent time}}_{\text{Channel 1b}} \\ &= \Delta\text{HH income} \times \text{Effect of HH income unit change} \\ &+ \Delta\text{Parent time} \times \text{Effect of parent time unit change} \end{aligned} \tag{5}$$

From Table 2, the 18-year-old's removal does not lead to higher mother earnings. Assuming based on this result that parents do not change the amount of time they spend on child care, the second term in equation (5)—the change in parent time spent on child care—is zero, and all of the effect of household resources loads onto the change in household income. The household loses $9,800 of SSI income when the younger sibling is less than 18 years old as a result of the 18-year-old's removal (see Table 2). Since the household resources effect is $3,300, this suggests that each additional $1,000 in total (not annual) household income during childhood increases annual adult earnings by $340.

5 Conclusion

Few programs have generated as much debate as the Supplemental Security Income program for children. Over the past decade, the voices in this debate have ranged from calls to defund the program due to perverse incentives to calls to increase the generosity of the program and expand the resources available to children with disabilities and their families. Missing from the debate has been empirical evidence on the ways in which SSI affects the human capital development and long-term outcomes of children. Using a pair of regression discontinuity designs and a novel matching procedure in SSA administrative data, this paper provides some of the first evidence on this question.

From the first RD design, I estimate that removing children from SSI before the age of 12 has a net effect on adult earnings that is not statistically distinguishable from zero. I expect this estimate to become more precise with additional years of earnings data. To identify the part of the net effect operating through household resources, I develop the first probabilistic matching procedure to identify siblings in SSA administrative data who do not themselves receive SSI benefits. From the second RD design, I find that the younger siblings of 18-year-olds who are removed from SSI experience large and persistent declines in earnings in adulthood, on the order of $5,000 annually from a baseline of $14,000.

The finding that removing children from SSI at age 18 has large and persistent adverse effects on the adult earnings of their younger siblings has important policy implications. Notably, this effect is not limited to the sample of siblings who themselves receive SSI in childhood, for whom the 18-year-old's removal might affect expectations about their own SSI receipt in adulthood. The effect is even stronger in the sample of siblings who do not receive SSI as children, suggesting that the effects on younger siblings are attributable to a decline in household resources that SSI was previously providing. This evidence suggests, for the first time, that SSI affects children's human capital development through household resources. With additional years of data, this analysis will also provide evidence on the magnitude of the perverse incentives effect relative to the household resources effect.

References

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney, "The long-run impact of cash transfers to poor families," *American Economic Review*, 2016, *106* (4), 935–71.

Coe, Norma B and Matthew S Rutledge, "What is the long-term impact of Zebley on adult and child outcomes?," *Center for Retirement Research at Boston College Working Paper*, 2013, *3*.

Deshpande, Manasi, "Does Welfare Inhibit Success? The Long-Term Effects of Removing Low-Income Youth from the Disability Rolls," *American Economic Review*, November 2016, *106* (11), 3300–3330.

_ , "The Effect of Disability Payments on Household Earnings and Income: Evidence from the SSI Children's Program," *The Review of Economics and Statistics*, 2016, *98* (4), 638–654.

Levere, Michael, "The labor market consequences of receiving disability benefits during childhood," *Journal of Human Resources*, 2019, pp. 1118–9883R1.

Price, David and Jae Song, "The Long-Term Effects of Cash Assistance," *Working Paper*, 2020.

Appendix (for online publication)

A Probabilistic sibling matching procedure

This appendix section describes the matching procedure for the construction of sibling Sample 2 in Section 4.2. The original sample for the Age-18 RD comprises 81,800 individuals from the Supplemental Security Record (SSR) who received SSI on their 18th birthday and whose 18th birthday was within 37 weeks of August 22, 1996, the date of PRWORA enactment. The SSR includes identifiers for these individuals, including their Social Security Number (SSN); first, middle, and last names; and date of birth. The SSR also includes identifiers for the parents of children (below 18 years) receiving SSI benefits, including parent SSN; parent first, middle, and last names; parent sex; and parent date of birth. Note that for the purposes of the SSR a "parent" need not be a biological parent, but is generally someone who lives with the child and has custody of the child. A child's parents in the SSR can change over time as a result of changes in family structure (e.g., divorce or remarriage). The children in the original Age-18 RD sample can be linked using SSN to their record(s) in the Numident, which includes first, middle, and last names at the time of birth and legal changes thereafter; date of birth; city and state (or country, if outside United States) of birth; citizenship; district office at which the Social Security card was processed; race; and mother's full name and father's full name (where available) at the time of the child's birth.

The goal of the matching procedure is to identify siblings of the 18-year-olds in the original Age-18 RD sample using the Numident. The challenge is that the Numident does not include unique identifiers for parents, only full mother and father name at the time of the child's birth. Many parent names are common names, and parent names (especially mother names) can change in between sibling births. I therefore develop a probabilistic matching procedure to identify likely younger siblings in the Numident. Since the 18-year-olds in the original Age-18 RD sample are born in 1978, I start with an extract of the Numident provided by the Office of Data Development (Office of Research, Evaluation, and Statistics) at SSA that includes all records for individuals in the United States with a date of birth between 1976 and 1990 (176 million records). I link the individuals in the original Age-18 RD sample to all mother and father names associated with that individual in either the SSR or the Numident. I then link these mother and father names to mother and father names in the entire Numident extract using exact name matching; Part 1 merges use first, last, and middle names, while Part 2 merges use only first and last names. I then link the individuals in the original Age-18 RD sample to all individuals in the Numident that share a parent name. Of course, this parent name matching procedure will include a huge number of false positives

due to common parent names.[12] To filter out false positives, I drop the following potential sibling matches:

- Potential Numident sibling's date of birth is less than 300 days apart or more than 12 years apart from the date of birth of the individual from the original Age-18 RD sample
- Potential Numident sibling's date of birth is less than 16 years or more than 40 years from the date of birth of the mother of the individual from the original Age-18 RD sample (known from the SSR)

To facilitate the probabilistic matching procedure, I create variables for the following characteristics of sibling matches:

- Siblings connected through one vs. two parents, and through parent first and last names only vs. first, last, and middle names
- Siblings share a city and/or state of birth
- Siblings have same district office listed on record
- Siblings have the same race
- Siblings have the same last name
- Annual earnings of the SSR mother and father of the 18-year-old relative to the potential Numident sibling's year of birth
- Frequency of parent combined first, middle, and last name of potential siblings

Using these characteristics, I conduct two exercises to determine which potential sibling matches are most likely to be true matches:

1. Predicting based on known siblings. Recall that a subset of siblings are known from the SSR (Sample 1)—these are siblings that also receive SSI benefits during childhood and share a parent SSN with an 18-year-old in the original Age-18 RD sample. Using these known sibling matches, I regress an indicator for known sibling match on the characteristics listed above and then predict the likelihood of being a sibling based on that regression. I keep the potential sibling matches with the top 5 percent of predicted scores.

[12]This matching procedure will also exclude many true matches because of changes in parent names. Although this issue is concerning for the precision and representativeness of Sample 2, it is less concerning from a bias perspective.

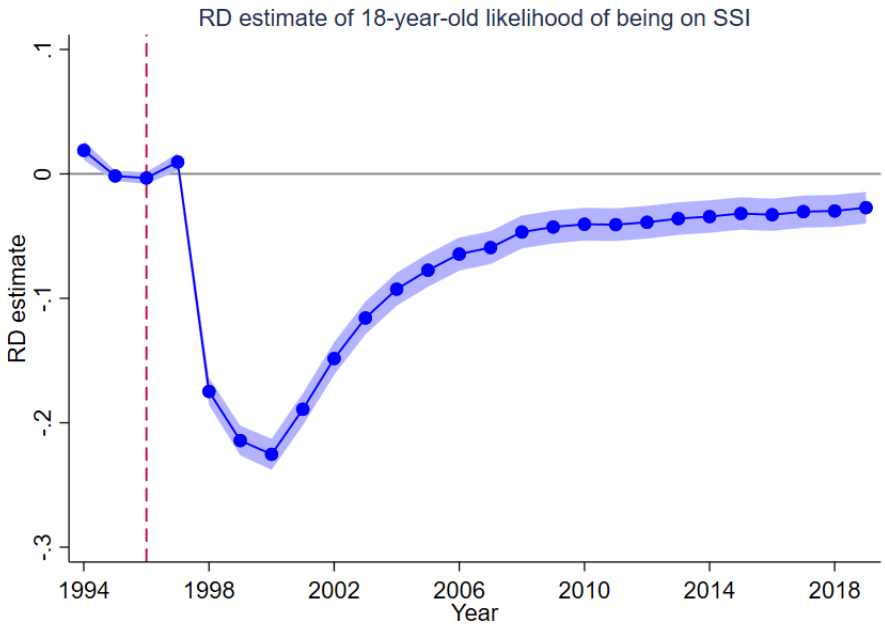
2. Tagging likely matches. SSR siblings may not be representative of the broader set of siblings that do not also receive SSI benefits. For this reason, I tag matches that are intuitively most likely to be true matches:

    - Siblings connected by both a mother name and father name, at least one of which uses parent middle name/initial
    - Siblings connected by a rare parent name (fewer than 5 of the first, middle, and last combination in the Numident extract), at least one parent middle name/initial, and same state of birth
    - Siblings connected by a rare parent name (see above) and the same city of birth

From these two samples, I further restrict to matches that share a state, city, or district office at birth; AND that are share two parents, one parent with a middle name match, or a last name. This leaves 4.1 million potential siblings. I also drop siblings who appear in Sample 1 and siblings that are born fewer than four years after the 18-year-old. This leaves 70,300 potential siblings.

Appendix Figures and Tables (for online publication)

Figure 7: Age-18 RD: First stage estimate over time



Notes: Figure plots the RD estimate (β from equation (3) of the likelihood that the 18-year-olds in the original sample are on SSI from 1994 to 2019. Vertical line at 1996 indicates the year of PRWORA enactment when the reform to age 18 medical reviews occurred.