The Long-Run Effects of Parental Unemployment in Childhood

James Uguccioni*

Latest Draft Available Here

November 9, 2021

Abstract

Parental job loss is a large, negative shock to the household that can affect children in both the short- and long-run. Little is known, however, about how the long-run impacts of job loss on children vary with the child's age at the time of displacement. This paper provides the first empirical evidence of the long-run effects of parental unemployment on children exposed before age 10 (and as young as 2), a period thought to be critical for child development. Using administrative tax data covering the universe of children born in Canada between 1972 and 1985 and random forest proximity matching, I estimate the causal effects of parental job loss experienced at different points in childhood on a child's income attainment. I find that children exposed to parental unemployment at ages 2 to 10 experience losses of 3 to 4 rank points in average earnings attainment in adulthood (approximately $2,500 per year). These children are also 36% more likely to receive welfare as adults and 4% less likely to pursue post-secondary education. Consistent with critical periods of child development, children who experience parental job loss before age 10 experience larger reductions in income attainment than children exposed at older ages. Decomposing these estimates, I show that the majority of my treatment effects are attributable to the timing of income losses experienced during childhood, as well as unemployment-induced moves to neighourhoods with less opportunity.

*University of Toronto, james.uguccioni@mail.utoronto.ca. I would like to thank Shari Eli, Kory Kroft, and Michael Baker for their support and supervision throughout this project, Carolina Arteaga, Jean-William Laliberté, Adriana Lleras-Muney, Ismael Mourifié, Philip Oreopoulos, David Price, Heather Sarsons, Clémentine Van Effenterre, members of the 5060 reading group, and seminar participants at the University of Toronto for their invaluable feedback. All errors are my own. Much of the analysis in this paper was conducted at the Toronto Research Data Centre, using administrative data made available by Statistics Canada in the Canada Research Data Centre Network (CRDCN). The views expressed in the paper are mine alone, and do not necessarily reflect those of the CRDCN or Statistics Canada.

1 Introduction

An unemployment spell represents a large, adverse shock to a household. Parents who lose their jobs not only face scarred earnings for years to come (Davis and Von Wachter, 2011; Stepner, 2019), but also increased incidences of anxiety and depression (Burgard et al., 2007; Brand and Burgard, 2008; Brand, 2015), opioid use (Morthorst et al., 2021), and divorce (Charles and Stephens, 2004). Children who experience parental unemployment have been shown to have lower self-esteem (Kalil and Ziol-Guest, 2005), worse educational attainment (Hilger, 2016; Tanndal and Paallysaho, 2020; Ruiz-Valenzuela, 2021), and lower income attainment in adulthood (Oreopoulos et al., 2008).

While some work has studied the effects of parental unemployment on children's outcomes when they are as young as 4 (Mork et al., 2019), data limitations have constrained research on children's long-run outcomes to children exposed at age 10 or older (Oreopoulos et al., 2008; Hilger, 2016; Tanndal and Paallysaho, 2020). Given that related work studying how neighbourhood quality affects children has found that moving to a better neighbourhood at a younger age benefits children more in the long-run (Chetty et al., 2016; Chetty and Hendren, 2018a,b; Chyn, 2018), it seems unlikely that estimates for these older children will generalize to their younger peers.

I provide the first empirical evidence of the long-run effects of parental job loss on children exposed at ages as young as 2. I employ detailed administrative tax data on parents and children to study how the timing of parental unemployment throughout childhood affects children in the long-run. My data cover the universe of children born in Canada between 1972 and 1985, linked to both their own and their parents' annual tax returns from 1981 to 2016. The scope of these data allows me to follow children and their households throughout childhood, and then relate their experiences to their income attainment in adulthood.

While parental unemployment is often an unexpected shock for the household, it is not randomly assigned throughout the income distribution. As such, a naïve comparison of children who did and did not experience parental unemployment would yield biased estimates of its effect on their income attainment. To address this identification challenge, I leverage the richness of my data and employ a state-of-the-art matching research design which builds on a selection on observables assumption. Specifically, I use a random forest to predict treatment propensity and match observations which are frequently sorted into the same leaves, as measured by the proximity matrix. My procedure ensures that matches are created along signal dimensions of the data, and allows me to exploit the rich data at my disposal without incurring the curse of dimensionality.

I estimate the effect of exposure to parental unemployment on a child's income attainment for children who were first exposed at the ages of 2 to 21, respectively. I find that children exposed to parental unemployment before age 10 experience larger losses in income attainment in adulthood than those exposed later in their childhoods. In particular, parental unemployment experienced at ages 2 to 10 reduces a child's income attainment by 2 to 4 percentiles in their birth cohort's income distribution. Children exposed after age 10 experience smaller losses in earnings attainment, with null effects for children exposed in their late teens. Alongside their reduced income attainment, I find that children who experience parental unemployment between the ages of 2 and 13 are 3% to 4% less likely to pursue post-secondary education and approximately 36% more likely to receive Social Assistance payments in adulthood.

To understand the dynamics among my estimated effects, I consider whether larger losses in income attainment for children exposed at younger ages are the result of greater exposure to the unemployment shock or the timing of the shock during childhood. As parental unemployment affects the household for many years after the initial job loss, children who are treated at younger ages will be exposed to its aftershocks for a larger fraction of their childhood and therefore may be more affected in the long-run. Alternatively, early childhood is understood to be a critical period for child development (Currie and Almond, 2011; Almond et al., 2018), so children treated at younger ages may miss out on higher returns on human capital investments. I find that lost parental earnings from unemployment are far more important for children exposed at younger ages, consistent with high returns on human capital investments made during critical periods of childhood.

As parental unemployment is a multifaceted shock to the household, I also consider whether children are affected through channels other than reduced household income. I show that unemployment causes households to move to neighbourhoods with lower educational attainment, as well as an increased incidence of divorce. Using a mediation analysis (Heckman et al., 2013), I find that parental income scarring accounts for 30% to 50% of my estimated treatment effects, while neighbourhoods explain an additional 4% of the estimated effects. Consistent with the literature (Corak, 2001), I find little evidence that experiencing divorce at any age affects a child's income attainment.

Overall, children who are exposed to parental unemployment at younger ages earn less in adulthood, are less likely to attend post-secondary education, and are more likely to receive social insurance benefits from the state. These effects are quantitatively large when compared with estimates of the intergenerational transmission of income. In Canada, Connolly et al. (2019) estimate that a 10 percentile rank increase in a parent's income is associated with a 2.1 percentile rank increase in their child's income. In my preferred specification, children exposed to parental unemployment at ages 2 to 3 experience a 3.7 percentile rank reduction in their income attainment at ages 27 to 31. The Connolly et al. (2019) rank-rank slope implies experiencing parental unemployment at age 2 or 3 is equivalent[1] to moving a child from a median income household to one in the 32$^{\text{nd}}$ percentile — a 17.6 percentile drop[2].

Alternatively, we can compare the intergenerational effects of unemployment with income losses resulting from a child's own layoff. Studying Canadian layoffs between 2003 and 2010, just before many of the children in my sample reached age 30, Stepner (2019) estimates that layoffs cause an average reduction in after tax income of 10% ($8,000) over five years. I find that experiencing parental unemployment at ages 6 or 7 causes a 6% ($2,500) reduction in average earnings attainment between the ages of 30 and 34[3]. As such, parental unemployment experienced early in childhood causes a reduction in earnings attainment roughly two thirds the size of the income losses caused by the child *themselves* being laid off.

[1]Of course, we have extensive evidence from adoption studies that the genetically inherited traits from parents explain a significant fraction of a child's educational attainment (Sacerdote, 2007) and wealth accumulation (Fagereng et al., 2021) in adulthood. As we expect that these traits are positively correlated with income for parents and children, the rank-rank slope likely overstates the effects of parental income on children. If the rank-rank slope overestimates the causal effect of parental earnings on children, then it will imply a lower bound on the reduction in household earnings equivalent to experiencing parental unemployment.

[2]These same children have parents earning roughly 9.3 percentiles below parents in the control group when the rank-rank slope is conventionally measured (parental earnings when the child is age 16 to 18). Of course, parental unemployment may affect children beyond income scarring, so we should expect that the reduction in parental earnings implied by the treatment effect is far larger than the realized loss.

[3]The differences between my estimates and the average income lost due to layoffs in Stepner (2019) reflects differences in sample construction. While my estimates only include individuals ages 30 to 34, Stepner (2019) includes individuals ages 25 to 55. Kaila et al. (2021) provides an alternative point of comparison, focusing on individuals who experience layoffs between ages 30 and 40 in Finland. They find that children whose parents had earnings in the bottom 20% of the income distribution, and are laid off themselves, experience a reduction in their earnings equivalent to 7 percentile ranks. To the extent those estimates generalize to Canada, my estimates imply that experiencing parental unemployment at ages 6 or 7 caused children to lose about two thirds as much income as they would if they were laid off themselves.

1.1 Related Literature

My paper contributes to three academic literatures. First and foremost, I add empirical evidence of the long-run effects of parental job loss on young children to the aforementioned literature studying how parental unemployment affects children (Oreopoulos et al., 2008; Hilger, 2016; Mork et al., 2019; Tanndal and Paallysaho, 2020). In particular, I show that the children exposed before age 10 experienced larger reductions in income attainment than children who were treated at the older ages. Building on previous work documenting how income losses (Oreopoulos et al., 2008) and household stress (Tanndal and Paallysaho, 2020) explain some of these long-run effects, I also present evidence that moves to neighbourhoods with lower educational attainment explain an important part of how parental unemployment affects children.

My paper also contributes to a literature studying critical periods in childhood when human capital investments may have particularly large returns (Currie and Almond, 2011; Almond et al., 2018). While there is ample empirical evidence of critical periods in the short-run (e.g. Doyle, 2020), this area remains underexplored when relating them to a child's well-being in the long-run. I add to context to previous work documenting the long-run benefits of social insurance programs, such as foodstamps (Hoynes et al., 2016) or cash transfers (Aizer et al., 2016), for children in need during critical periods of their lives by focussing on the target of another major social insurance program — unemployment insurance. I find evidence that parental unemployment has large effects on children before age 5, commonly understood to be the most important critical period for a child's development (Currie and Almond, 2011). Unlike much of the literature on critical periods, I do not find that effects are smaller for children exposed after age 5. Rather, I find evidence of similarly large effects of parental unemployment experienced between the ages of 7 to 10.

Finally, I provide evidence that the timing of parental earnings can affect a child's income attainment, adding to the literature studying intergenerational income transmission (Becker and Tomes, 1979, 1986; Solon, 1999; Chetty et al., 2014). Empirical work in this literature commonly treats childhood as a single period, and uses a parent's income over a handful of years (e.g. Chetty et al., 2014) to study its relationship with their child's income attainment. Of course, if parents are not credit constrained, there is no need to distinguish between a snapshot of childhood and its entirety. In some cases, ignoring these intragenerational credit constraints has been empirically validated[4]. Other simulated (Caucutt et al., 2017; Caucutt and Lochner, 2020) and empirical (Carneiro et al., 2021) work, however, has stressed that the timing of a parent's earnings can affect a child's income attainment if intragenerational credit constraints bind. I find that the importance of parental income is decreasing in the age at which a child is exposed to parental unemployment, consistent with parents who are borrowing constrained. Previous work in this literature has looked at a family's inability to borrow against the future earnings of children as a credit constraint which increases intergenerational income inequality (e.g. Herbst and Hendren, 2021). My paper suggests that a parent's inability to borrow against their own future earnings could be an additional source of intergenerational inequality.

2 Empirical Model

Parental unemployment is a complex shock for children to experience. Income scarring and weakened labour force attachment after the initial job loss both mean that children are affected for years after they are first exposed. As a result of these dynamics, experiencing parental unemployment at age a is very different than a hypothetical randomized control trial which randomly takes cash from families when a child is age a.

To fix ideas, I present a simple potential outcomes framework where childhood lasts two periods and parents can experience unemployment in either period[5]. Using these potential outcomes, I establish identification of my treatment effects of interest under a conditional independence assumption and map target parameters to my estimand. This Section informs the interpretation of the estimates presented in Section 5. Readers who are strictly interested in my empirical results may wish to proceed to Section 3, and can do so without hampering their understanding of the paper.

2.1 Defining Target Parameters

Suppose a parent lives for 2 periods, and is endowed with a child which she raises during this time. Parents are uncertain of their employment status $d_a \in \{0, 1\}$ in each period when their

[4]For example, work studying neighbourhood quality and its effects on a child's income attainment (Chetty and Hendren, 2018a,b) has shown that the age of exposure to a better neighbourhood does not matter so much as the length of exposure to one, consistent with parents who are not borrowing constrained.

[5]In Appendix A1, I present an economic model where parents make human capital investments in their children and are uncertain of their own future employment status. These potential outcomes arise directly from this model.

child is age $a \in \{1, 2\}$. At each age a, parents are learn whether they are employed, $D_a = 0$, or unemployed, $D_a = 1$, and then make investments in their child's human capital. Using the same notation as Torgovitsky (2019), let $\{U_a(0), U_a(1)\}$ denote the potential outcomes for employment D_a when $D_{a-1} = 0$ and $D_{a-1} = 1$ respectively. In particular, a parent's realized employment status is given by,

$$D_a = (1 - D_{a-1})U_a(0) + D_{a-1}U_a(1)$$

Given an initial condition, D_0, and the set of all potential outcomes, $\{U_a(0), U_a(1)\}_{a=1,2}$, we can fully describe a parent's employment history over these two periods. We can also relate the various possible employment histories to my observed outcome of interest, Y, the child's income attainment at ages 27 to 31. In particular, let $Y(d_1, d_2)$ be the income attainment of a child when parental employment history is (d_1, d_2). Realized income attainment for the child is given by,

$$Y = \sum_{d_1 \in \{0,1\}} \sum_{d_2 \in \{0,1\}} Y(d_1, d_2)\mathbf{1}\{D_1 = d_1\}\mathbf{1}\{D_2 = d_2\}$$

In this paper, I study how the timing of a parent's first unemployment spell affects a child's income attainment. In this simple framework, we can express how parental unemployment affects children who are first exposed at age 1 or first exposed at age 2 (i.e. they were not first exposed at age 1). For children who are first exposed at age 1, $D_1 = 1$, it is possible for their parents to either be employed or unemployed when they are age 2. As such, children exposed at age 1 will realize either $Y(1, 0)$ or $Y(1, 1)$ depending on their parent's potential outcome $U_2(1)$. Similarly, children who are not first exposed at age 1 will realized either $Y(0, 0)$ or $Y(0, 1)$ depending on their parent's potential outcome $U_2(0)$. As a result, there are four potential treatment effects of interest depending on a parent's potential employment outcomes, $(U_2(0), U_2(1))$,

$$Y(1, 1) - Y(0, 0)$$

$$Y(1, 1) - Y(0, 1)$$

$$Y(1, 0) - Y(0, 0)$$

$$Y(1, 0) - Y(0, 1)$$

The potential outcomes for children first exposed to parental unemployment at age 2 are considerably more straightforward. Children who are first exposed at age 2 necessarily realize income attainment $Y = Y(0,1)$. Children who are not first exposed to parental unemployment at age 2, the terminal period of their childhood, will never be exposed to parental unemployment and realize income attainment $Y = Y(0,0)$. As a result, there is only one potential treatment effect of interest which does not depend on the parent's potential employment outcomes,

$$Y(0,1) - Y(0,0)$$

2.2 Identifying Target Parameters

Estimating a treatment effect such as $Y(0,1) - Y(0,0)$ presents considerable identification challenges. In particular, the unemployed tend to be negatively selected (Davis and Von Wachter, 2011), and a parent's innate abilities are inherited by their children (Sacerdote, 2007; Fagereng et al., 2021). If an unemployed parent is negatively selected, then we should expect that her child will be negatively selected among all children as well. In this setting, a naïve comparison of the average income attainment for children who first experience parental unemployment at age 2 with those who did not first experience it at age 2 would likely produce an estimate of the causal effect $Y(0,1) - Y(0,0)$ which is negatively biased. That is,

$$\begin{aligned} & \mathbb{E}[Y|D_1 = 0, D_2 = 1] - \mathbb{E}[Y|D_1 = 0, D_2 = 0] \\ = \; & \underbrace{\mathbb{E}[Y(0,1) - Y(0,0)|D_1 = 0, D_2 = 1]}_{\text{Causal Effect}} \\ & + \underbrace{\mathbb{E}[Y(0,0)|D_1 = 0, D_2 = 1] - \mathbb{E}[Y(0,0)|D_1 = 0, D_2 = 0]}_{\text{Selection into Treatment}} \qquad (1) \end{aligned}$$

where we expect $\mathbb{E}[Y(0,0)|D_1 = 0, D_2 = 0]$ is higher than $\mathbb{E}[Y(0,0)|D_1 = 0, D_2 = 1]$ due to the negative selection of parents into unemployment. I identify the causal effect of first experiencing parental unemployment at age a under a selection on observables assumption, conditioning on vector of observables $x \in \mathbf{X}$ which I assume are exogenous to treatment.

A1 Common Support For every child who is first treated at age $a \in \{1, 2\}$ with observables $x \in \mathbf{X}$, there is a non-zero probability of not being treated.

$$\textbf{A1.1} \qquad Pr(D_1 = 1|X = x) \in [0, 1) \;\; \forall x \in \mathbf{X}$$

$$\textbf{A1.2} \qquad Pr(D_2 = 1|X = x, D_1 = 0) \in [0, 1) \;\; \forall x \in \mathbf{X}$$

We can equally rephrase **A1** in terms of the parent's potential employment outcomes $\{U_a(0), U_a(1)\}$. In this case, **A1.1** is equivalent to assuming that there is no point $x \in X$ such that all parents will be unemployed with certainty when their child is age 1, $U_1(1) = U_1(0) = 1$[6]. A similar extension applies to **A1.2**, however children who may first experience parental unemployment at age 2 did not at age 1 by definition. As we are conditioning on $D_1 = 0$, we only need to assume that there is no point $x \in X$ such that all parents who are employed in period 1 will lose their jobs with certainty, $U_2(0) = 1$.

A2 Conditional Independence of Potential Outcomes First experiencing treatment at age $a \in \{1, 2\}$ is as good as randomly assigned conditional on observables X.

$$\textbf{A2.1} \qquad \{Y(0,1), Y(0,0), U_2(0), U_2(1)\} \perp\!\!\!\perp D_1 | X = x \;\; \forall x \in \mathbf{X}$$

$$\textbf{A2.2} \qquad \{Y(0,0)\} \perp\!\!\!\perp D_2 | X = x, D_1 = 0 \;\; \forall x \in \mathbf{X}$$

As with common support, the interpretation of the conditional independence assumption **A2** varies slightly depending on whether children are treated in a non-terminal period of their childhood or in the terminal period. Children first treated at age 1 are treated in a non-terminal period of their childhoods, and hence their treatment status must be conditionally independent of all counterfactual outcomes they could have experienced, $\{Y(0,0), Y(0,1)\}$.

As there is more than one counterfactual outcome these children could have realized, it is also important that their parent's potential employment outcomes $\{U_2(0), U_2(1)\}$ are conditionally independent of treatment. These potential employment outcomes help us model which the two potential income attainment outcomes a treated child would have realized if her parents were not unemployment in period 1. It is helpful to sort parents into the four pairs of $\{U_2(0), U_2(1)\}$ which could be realized:

[6]Specifically, $\mathbf{A1.1'}$ $Pr(U_1(1) = U_1(0) = 1|X = x) \in [0, 1)$ $\forall x \in \mathbf{X}$. For **A1.2**, we have $\mathbf{A1.2'}$ $Pr(U_2(0) = 1|X = x, D_1 = 0) \in [0, 1)$ $\forall x \in \mathbf{X}$

	$U_2(0) = 0$	$U_2(0) = 1$
$U_2(1) = 0$	Always employed (E)	Negative state dependence (NS)
$U_2(1) = 1$	Positive state dependence (PS)	Always unemployed (U)

Children whose parents exhibit positive state dependence, for example, will realize $Y(1,1)$ if their parents experience $D_1 = 1$. If they had counterfactually not experienced parental unemployment at age 1, $D_1 = 0$, then their parents would have also been employed in period 2 and they would have attained income $Y(0,0)$. For these children, the lost income attainment which results from first experiencing parental unemployment in period 1 is $Y(1,1) - Y(0,0)$. Crucially, notice that assuming $\{U_2(0), U_2(1)\} \perp\!\!\!\perp D_1|X$ is a statement on potential outcomes, and it does not imply $D_2 \perp\!\!\!\perp D_1|X$. Under **A2.1**, parents with $U_2(1) = 0$, for example, are no more likely to be unemployed in period 1 than they are to be employed after conditioning on X.

Assumptions **A1** and **A2** solve the negative selection into treatment presented in equation (1). Under Assumption **A2.1**,

$$\mathbb{E}[Y(0,0)|D_1 = 0, D_2 = 1, X = x] = \mathbb{E}[Y(0,0)|D_1 = 0, D_2 = 0, X = x]$$

and Assupmtion **A1.1** then ensures that this applies at every point of support $x \in \mathbf{X}$ where $Pr(D_2 = 1|X = x, D_1 = 0) \in [0,1)$. Similarly, under Assumption **A2.2** for $d_2, d_2' \in \{0,1\}$,

$$\begin{aligned} & \mathbb{E}[Y(0,d_2)|D_1 = 1, U_2(0) = d_2, U_2(1) = d_2', X = x] \\ = \quad & \mathbb{E}[Y(0,d_2)|D_1 = 0, U_2(0) = d_2, U_2(1) = d_2', X = x] \end{aligned}$$

solving the selection into treatment issue at a particular x. Assumption **A1.2** then ensures that this applies at every point of support $x \in \mathbf{X}$ where $Pr(D_1 = 1|X = x) \in [0,1)$.

2.3 Estimating Target Parameters

I use a matching research design to estimate the effect of first experience parental unemployment at age a on a child's income attainment. I create a matched sample of children first treated at age a linked to counterfactuals with similar observables $x \in \mathbf{X}$ who were not first treated at age a. I discuss my matching strategy at length in Section 4.1. Using this matched sample, I estimate the average reduction in income experienced by a child first exposed to parental unemployment at age a,

$$\mathbb{E}_X[\mathbb{E}(Y|D_1 = 1, X = x)] - \mathbb{E}_X[\mathbb{E}(Y|D_1 = 0, X = x)] \tag{2}$$

$$\mathbb{E}_X[\mathbb{E}(Y|D_1 = 0, D_2 = 1, X = x)] - \mathbb{E}_X[\mathbb{E}(Y|D_1 = 0, D_2 = 0, X = x)] \tag{3}$$

Under assumptions **A1** and **A2**, my estimates are best interpreted as conditional average treatment effects on the treated (CATTs). They are conditional as I fix the age of first exposure to parental unemployment, and treatment on the treated as unemployment is extremely rare for some values of $x \in \mathbf{X}$[7].

The causal effect of experiencing parental unemployment at age 1, presented in equation (2), is a weighted average of the four possible treatment effects presented in Section 2.1. Each of these treatment effects is weighted according to the probability of observing a parent with potential outcomes $(U_2(0), U_2(1))$ at every $x \in \mathbf{X}$[8]. In particular,

$$\begin{aligned}
& \mathbb{E}_X[\mathbb{E}(Y|D_1 = 1, X = x)] - \mathbb{E}_X[\mathbb{E}(Y|D_1 = 0, X = x)] \\
= \; & \mathbb{E}_X\left(Pr(X = x, PS)\, \mathbb{E}\left[Y(1,1) - Y(0,0)|D_1 = 1, X = x, PS\right]\right) \\
& + \mathbb{E}_X\left(Pr(X = x, NS)\, \mathbb{E}\left[Y(1,0) - Y(0,1)|D_1 = 1, X = x, NS\right]\right) \\
& + \mathbb{E}_X\left(Pr(X = x, U)\, \mathbb{E}\left[Y(1,1) - Y(0,1)|D_1 = 1, X = x, U\right]\right) \\
& + \mathbb{E}_X\left(Pr(X = x, A)\, \mathbb{E}\left[Y(1,0) - Y(0,0)|D_1 = 1, X = x, A\right]\right)
\end{aligned}$$

It is helpful to notice the role played by **A2.1** relating my estimand to the possible treatment effects in the population. If treatment status is conditionally independent of a child's potential outcomes, then $Y(0,0)$ and $Y(0,1)$ approximate what would have happened to a treated child if she had experienced parental unemployment history $(0,0)$ or $(0,1)$ respectively. Similarly, conditional independence of $U_2(0)$ and $U_2(1)$ ensure that treated children are no more likely to have parents who exhibit state (in)dependence than counterfactual children.

Holding the weights on each group fixed, the roles played by the timing of and the exposure to parental unemployment in a child's income attainment become apparent. Parents in NS exhibit negative state dependence, $(U_2(0) = 1, U_2(1) = 0)$, and the scale of their child's reduced income attainment from experiencing parental unemployment at age 1 is given by

[7] For example, I will not observe how treatment would affect a child with two parents who are tenured professors, and that lack of common support does not allow me to estimate an average treatment effect.

[8] In Appendix A2 I formally establish this identification result.

$Y(1,0) - Y(0,1)$. Children with parents in NS will experience one year of exposure to parental unemployment in their lives, so their treatment effect will speak to the importance of human capital investments made at age 2 relative to age 1. If, for example, age 1 is a critical period for investments, then we should expect that $Y(1,0) < Y(0,1)$. Alternatively, if the timing of parental unemployment is relatively unimportant, then we should expect that $Y(1,0) \approx Y(0,1)$ and children with parents in NS will not contribute much to my estimated effect.

Parents in U and A do not exhibit state dependence, $U_2(0) = U_2(1)$, and the scale of their child's reduced income attainment from experiencing parental unemployment at age 1 governed by their total exposure to the job loss. For children with parents in U, first experiencing parental unemployment at age 1 means they will experience two years of unemployment rather than one. Similarly, for children with parents in A, first experiencing parental unemployment at age 1 means they will experience one year of unemployment rather than none. Their importance in my estimated effect will dependent on how much worse off an additional year of parental unemployment leaves the affected children.

Finally, children whose parents exhibit positive state dependence, PS, face both effects. If they are exposed to parental unemployment at age 1, these children will necessarily experience parental unemployment in the second period as well. As such, treatment at age 1 increases their exposure to parental unemployment by two years. If timing is important, and there is a critical period for human capital investments, these children will also experience parental unemployment during the critical period.

Age 2 is the terminal period of childhood, making the effect of experiencing parental unemployment at age 2 easier to interpret. As the parent's future employment status does not affect the child's income attainment, there is only one possible potential outcome for treated and counterfactual children respectively. Under assumptions **A1.2** and **A2.2**,

$$\begin{aligned} & \mathbb{E}_X[\mathbb{E}(Y|D_1 = 0, D_2 = 1, X = x)] - \mathbb{E}_X[\mathbb{E}(Y|D_1 = 0, D_2 = 0, X = x)] \\ = \; & \mathbb{E}\left[Y(0,1) - Y(0,0)|D_1 = 0, D_2 = 1\right] \end{aligned}$$

The intuitions from this model readily maps to my empirical application. In the data, I treat childhood as 21 periods (years) rather than two. Children who are first treated before age 21, the terminal period of childhood in my application, will experience income losses which reflect their parent's potential employment outcomes $\{U_a(0), U_a(1)\}_{a\in\{1,2,...,21\}}$. The number of different potential outcomes for child first exposed to parental unemployment at age a is increasing in amount of childhood remaining, and as such my estimand will average of more and more treatment effects for children treated at younger ages[9]. Focussing on the effect of first being exposed to parental unemployment at age a keeps this problem tractable, while still producing estimates with a clear interpretation.

3 Data

To study the effects of parental unemployment on children, I employ a rich administrative dataset, the Intergenerational Income Database (IID), of income tax returns in Canada from 1981 to 2016. These data cover approximately 3.6 million children born in 1972-1975, 1977-1980, and 1982-1985, accounting for around 77% of all Canadian children born over those years. The data were constructed by sampling all children aged 16 to 19 years old in a given linkage year, then linking any tax return the child filed while living with her parents to those filed by her parents in the linkage year or any of the four years thereafter[10].

The scope of the panel and the detailed administrative data provide several empirical advantages. Foremost among these advantages is the length of the IID. Any study of the effects of unemployment during childhood on income attainment in adulthood needs to follow children from a young age through to an age when the life cycle bias[11] in earnings plateaus — typically understood to be early- to mid-30s. As a result of these data demands, the literature has been restricted to studying children who experience parental unemployment at older ages and require shorter panels to study, such as children aged 10 or older in Oreopoulos et al. (2008). An alternative route to reduce demands on the length of the panel is to study educational outcomes observed in late teens or early 20s (Hilger, 2016; Tanndal and Paallysaho, 2020) as a proxy for income. The IID is much less restrictive than past

[9]For a child first exposed to parental unemployment at age a, there are 2^{21-a} potential outcomes which can be realized. Similarly, there are 2^{21-a} potential outcomes which can be realized by children who are not first exposed to parental unemployment at age a, resulting in $2^{2(21-a)}$ possible treatment effects of interest. This means there are over 274 billion possible treatment effects for children who are treated at age 2.

[10]For example, Statistics Canada defined the sample frame to be all children aged 16 to 19 in 1991, covering children born between 1972 and 1975. Any income tax returns filed by the children and parents at the same address in the window of 1991-1995 were then linked to define a family. Statistics Canada performed this procedure in 1991, 1996, and 2001, resulting in the 1976 and 1981 birth cohorts not being included in the sampling frame in any linkage years

[11]Research has shown that studying earnings of children early in their working lives will understate the degree of inequality in their permanent income, while studying them late in the their working lives will overstate inequality. The literature has dubbed this measurement issue the life cycle bias (for more discussion see Nybom and Stuhler, 2016).

datasets employed by the literature, covering 35 years of tax returns which follow children and their parents from when the child is age 9 at the absolute oldest to age 31 at the absolute youngest[12].

Income taxes in Canada are filed by the individual, unlike joint filing for married couples in countries such as the United States. Consequently, an observation in these data is best understood as a child i linked to a parent $j(i)$. Once a linkage is created, all tax returns filed by i or $j(i)$ between tax years 1981 and 2016 are uniquely identified and included in this dataset to create a panel. Despite income taxes being evaluated at the individual-level, the income tax return does require individuals to indicate their spouse[13], and failing to do so can lead to incorrect tax returns and associated penalties. The panel of tax returns from child i and parent $j(i)$ therefore identify their spouses in any given year, and these records are also included in the data. The final dataset is a panel of tax returns for child i, their parent(s) $j(i)$ at time of linkage, and any spouses of i or $j(i)$ who appear in tax years between 1981 and 2016.

Unlike surveys such as the Panel Study of Income Dynamics, my data do not suffer from attrition other than individuals either failing to declare taxable income or moving abroad. Administrative data can, however, suffer from sample selection issues which result from dataset design. Of particular concern are the 23% of children born in 1972-1975, 1977-1980, or 1982-1985 who are not matched to parents and therefore not included in the IID. If, for example, children from of wealthier households are less likely to work in their teens and less likely to file taxes, then they will be less likely to be matched to their parents and the IID will not be representative of the population at large. Internal work at Statistics Canada (Simard-Duplain and St-Denis, 2020) has shown that the children in the IID are somewhat positively selected from the whole population along parental education and are more likely to come from two-parent households. Both of these dimensions are consistent with households which are more like to include a tax filing parent in each year, and therefore more likely to be matched to their children. While similar trends have been shown in past academic work using the IID (Oreopoulos, 2003), this selection only affects the external validity of my

[12]The cohort born in 1972 is aged 9 at the time the first tax returns included in the IID are filed in 1981. All other birth cohorts are younger when their parents file their first tax returns that are included in the IID in 1981, with the four youngest cohorts (1982-1985) having childhoods completely covered by the IID. The cohort of children born in 1985 are aged 31 at the end of the panel in 2016, while the oldest cohort is 44.

[13]Some tax credits, deductions, and taxable benefits are evaluated using household income.

results and should not affect my identifying assumptions **A1** and **A2**.

3.1 Outcomes

Following the literature on intergenerational income mobility (Solon, 1999; Chetty et al., 2014; Connolly et al., 2019), my primary outcome is a child's income attainment in adulthood. Recent work in this space has employed percentile rank in the income distribution (Chetty et al., 2014; Chetty and Hendren, 2018a; Kaila et al., 2021) as the main outcome of interest, as income rank is generally understood to suffer from lifecycle bias less than log average income[14]. In light of this body of work, I employ rank of average earnings between 27 and 31 as my primary outcome of interest. Income ranks are computed within a birth cohort, using various income concepts which are relevant to the child's well-being: labour income (wages and salaries measured on the T4 form), total taxable income, total income net of taxes, and household income net of taxes[15]. For robustness, I consider rank of average earnings between ages 30 and 34 for children born before 1983 as well.

To capture a broader sense of the child's well-being in adulthood, I consider several alternative outcomes associated with socio-economic status. To proxy post-secondary attendance, I use an indicator for whether a child ever claimed tuition and education tax credits when they are aged 18 to 25[16]. Given the family economics literature documenting the income insurance benefits of marriage (e.g. Chiappori et al., 2018), I also consider whether the child is married or has a common-law partner at ages 25 and 30. To capture family structure, I make use of dependants appearing on the tax returns for tax credits[17], using age at first birth and number of children at age 30 as additional outcomes. Finally, I consider receipt of

[14]As income rank suffers less from the lifecycle bias, it means we can credibly measure a child's income attainment at younger ages that we would be able to with log average income. Consequently, authors employ different ages to measure the child's earnings in their main analysis varying from ages as young as 24 (Chetty and Hendren, 2018a) to 30 (Chetty et al., 2014). In a recent paper employing the IID, Connolly et al. (2021) uses the rank of average earnings between 30 and 36 as their outcome of interest to safeguard against the measurement error which comes with using annual earnings as a proxy for permanent income. For robustness, I consider log average income for each measure as well.

[15]Note that T4 income does not include self-employment income, however it is included in total income measures. Household income net of taxes is defined as the sum of the child and her spouse respective total income net of tax. Only spouses who are considered common-law or married are included as they are listed by the child on her tax return. Roommates or co-habitating partners who are not yet considered common-law will not be included in household income.

[16]Tuition tax credits only appear in the data in 1997, when children born in the early 1970s may have completed their undergraduate degrees. To ensure I do not miss students who graduate quickly, I consider post-secondary attendance between the ages of 18 and 20 as well. I present results for regressions focussing to children who were ages 18 in 1997 (i.e. born after 1978) and children who were age 20 in 1997 (i.e. born after 1976) to ensure all birth cohorts are observed at "core" post-secondary ages.

[17]Until 2002 for the Family Allowance credit, thereafter for Canada Child Tax Benefit, National Child Benefit, and Universal Child Care Benefit.

income supplements such as Social Assistance and Employment Insurance[18] in adulthood to capture whether parental unemployment affects a child's use of the social safety net.

In Section 6, I use my panel on households during childhood to study how parental unemployment can affect children. To capture losses in income for the unemployed parent, I use earnings reported on all tax returns of the unemployed parent. As my main outcome of interest is measured in percentiles, I follow the literature (Chetty et al., 2014) and measure the parent's using her income rank in her child's birth cohort's income distribution. Of course, a one percentile change in earnings can translate to losses of very different scales throughout the income distribution. I also consider log averages of income to capture the scale of earnings losses experienced by households after parental unemployment. Changes in a parent's marital status from year to year allows me to follow marriage and divorce trends following unemployment. To capture the importance of neighbourhood quality in a child's income attainment, I follow Connolly et al. (2019) and use 1996 Census Divisions as a Canadian analogue of Commuting Zones in the United States. I merge estimates of intergenerational income mobility at the Census Division-level from Connolly et al. (2019), as well as information on the average educational attainment of adulthoods in these areas, to the parent's location in each year to capture the quality of neighbourhood the children experience at a given age.

3.2 Treatment Status

The information included on tax returns provides a myriad of outcomes to study how parental unemployment affects children. The IID does not, however, include direct information from employers on layoffs or information from individuals surveying their employment status[19]. Fortunately, unemployment insurance (UI) benefits are taxable in Canada and are therefore reported on the tax form. I employ information on UI from tax returns to define treatment status in Section 4.1.

UI benefits during my period of interest were readily available to most workers in the

[18] Note that Unemployment Insurance in Canada was renamed Employment Insurance in 1996. As a result, when discussing the parents in my sample I often refer to Unemployment Insurance receipt, while I refer to Employment Insurance receipt for children. While there are important differences in the eligibility rules for these two program, both are meant to insure workers against unemployment.

[19] Other work on layoffs in Canada has linked individual tax returns to firm data on employees using the Record of Employment tax slips (Stepner, 2019) or the Longitudinal Worker File (Huynh et al., 2017). Following discussions with Statistics Canada, it became apparent that linking either of these datasets to the IID was not feasible.

event of an unemployment spell. Eligibility prior to 1996 was determined by the number of "eligible weeks" of work an individual had in the 52 weeks preceding a claim. The minimum number of eligible weeks required to claim UI varied from 10 to 20 depending on the regional unemployment rate, with higher unemployment regions facing lower minimum requirements. For a week to be considered eligible, workers had to meet a minimum weekly hours or a minimum weekly earnings threshold, with both set low enough that most part-time workers qualified[20]. Claimants satisfying the eligibility requirement were then entitled to UI benefits equal to 60% of earnings over the eligible weeks up to a maximum benefit period for up to 50 weeks.

In the United States, there is a large body of work documenting UI take-up as low as 45% among all eligible claimants (Lachowska et al., 2021). If the same institutional impediments to claiming UI benefits existed in Canada, then employing receipt of benefits as an indicator of treatment status could place unknown restrictions on the set of treated parents included in my analysis. Fortunately, UI take-up is significantly higher in Canada. In work studying UI benefit receipt in the early 1980s, Storer and Audenrode (1995) estimate that approximately 83% of all eligible claimants receive benefits, and attribute the remainder primarily to eligible beneficiaries who found jobs in the two week gap between a layoff and beginning to receive benefits. The structure of Canadian UI, particularly before it was reformed in 1995, creates incentives for firms to encourage unemployed workers to apply for UI (Nakamura and Diewert, 2000), and as such UI receipt is a much better tool to identify unemployment spells in Canada than it may be in other contexts[21].

I focus on parental unemployment experienced between 1987 and 1993 to ensure consistency in my definition of treatment. Prior to 1987, the eligibility requirements for UI were

[20]From 1987 to 1995, a week of work was "eligible" if the individual worked a minimum of 15 hours in a week or earned 20% of maximum benefits in a week (Lin, 1998). Though the maximum benefits vary, in 1995 the maximum weekly benefit was $448 (Gray, 2006) so the minimum weekly earnings to qualify for UI was $89.60. In 1995 the median, prime age worker earned around $25,900 in labour income (see Statistics Canada Table 11-10-0239-01) which amounts to approximately $498 in weekly earnings when spread across a full year, so the minimum weekly earnings for an eligible week was about 18% of a median worker's weekly earnings. While minimum wages vary provincially, a minimum wage worker in Ontario in 1995 would have to work just over 13 hours in a week (at $6.85 per hour) to achieve the minimum earnings threshold.

[21]In the United States, the literature has explained low UI take-up with relatively low earnings replacements (Anderson and Meyer, 1997), behavioural impediments (Vroman, 2009; Ebenstein and Stange, 2010), or incentives for firms to challenge UI claims and thereby discourage their filing altogether (Lachowska et al., 2021). Historically, the Canadian system has been criticized for being too generous in defining eligibility based on weeks of work rather than hours (Human Resources Development Canada, 1994), it has paid higher benefits than American counterparts (Baker et al., 1998), and it has created incentives for firms to *encourage* workers to apply for UI (Nakamura and Diewert, 2000). In particular, UI claims are not experience rated in Canada and layoffs do not affect the taxes faced by a firm as they do in the United States. Before 1996, eligibility for UI benefits was determined by a minimum number of weeks of work rather than hours of work. The result was many firms structuring seasonal contracts to ensure workers would be eligible for UI during the off-season in order to shift labour costs from the firm to the government (Human Resources Development Canada, 1994).

somewhat more stringent[22] and therefore the inclusion of earlier unemployment spells may reflect a different group of parents. In 1996, the *Employment Insurance Act* reformed entire the UI system, and as such 1995 is the final year with the eligibility rules from 1987 in place. I employ 1993 as my preferred final period to avoid any contamination from responses to tax reforms in 1994 (Lavecchia and Tazhitdinova, 2021) and the reforms of the *Employment Insurance Act* of 1996 made public as recommendations in 1994 (Human Resources Development Canada, 1994). Ultimately, in my preferred approach I observe children who experienced parental unemployment between the ages 2 and 21. The number of years which I observe children at each of ages varies systematically with the birth cohorts included the IID. For example, I only observe one birth cohort at age 2 (the 1985 cohort in 1987) but I observe 7 birth cohorts at age 10 (one every year except 1991).

In Table 1, I compare fathers in the IID with fathers the public use data from the Labour Force Survey (LFS). Consistent with Simard-Duplain and St-Denis (2020), fathers in the IID are more like to have spouses in the labour force than fathers in the LFS. Regardless of whether a father reports being unemployed or employed in the LFS, they are more likely to report being married than fathers in the IID. Fathers in the IID also have more dependants than fathers in the LFS. In both datasets, fathers in the Atlantic provinces make up a much larger fraction of the unemployed than they do of the employed — a well known artefact of UI and the seasonal fishing industry. Fathers in Québec as also somewhat overrepresented among the unemployed population relative to the employed, while the opposite is true of fathers in Ontario or the Prairie provinces. Given the similarity between the unemployed parents in the LFS and the IID, my estimates are best understood as the effect of treatment on the treated.

4 Empirical Strategy

Under **A1** and **A2**, I can identify the effect of first experiencing parental unemployment at age a on a child's income attainment. In the data, I estimate these treatment effects by matching each child i first treated at age a with a counterfactual child i' who is not first treated at a. I create these matches by first using a random forest to predict a child's

[22] In particular, prior to 1987 an "eligible week" was defined by an individual who worked a minimum of 15 hours in a week and earned 20% of maximum benefits (Lin, 1998). In 1987, the eligibility changed to an individual who worked a minimum of 15 hours in a week *or* earned 20% of maximum benefits.

treatment status at age a for each birth cohort. I then match treated and control children who are frequently sorted into the same leaves in the decision trees which make up my forest, using the out-of-bag proximity matrix.

Using this matched sample, I estimate the effect of first experiencing parental unemployment at age a, Δ_a, on a child's income attainment, Y_i. In particular, I run regressions of the form,

$$Y_i = \lambda_a + \Delta_a D_i(a) + x_i'\gamma_a + \epsilon_i \tag{4}$$

where $D_i(a)$ is a treatment indicator equal to 1 if a child is first treated at a and equal to 0 if a child is a matched countefactual for treatment at age a. In equation (4), λ_a allows for differences in the average earnings for the matched counterfactuals corresponding to children first treated at age a. I also include child-level controls, x_i, which were not restricted in the matching procedure and are independent of treatment, but may affect earnings, such as the child's sex.

In my preferred specifications, the only control I include is the child's sex and I restrict $\gamma_a = \gamma\ \forall a$ to maximize statistical power, though the results are qualitatively robust to alternative specifications. Crucially, notice that my estimating equation (4) does not have any time variation in the child's outcome, and as such does not permit individual level fixed effects. I rely on the quality of my matches to ensure that there are no systematic differences in a child's time invariant heterogeneity between treatment and control groups.

Following similar work in the literature (Stepner, 2019), I match children using observables on their parents, $j(i)$ and $j(i')$, from year $t-3$ before their father's unemployment spell in year t. To test whether the matched sample created by my procedure satisfies my identifying assumptions, I investigate whether there are any observable differences between treated and control households in all years $\tau < t-3$. In the absence of any clear differences, I argue that estimating my treatment effects of interest using the matched sample will satisfy my identifying assumptions.

4.1 Proximity Matching

Assumptions **A1** and **A2** do not readily apply to the data. In the case of continuous matching variables, the probability of two observations realizing the exact same value of a given variable is infinitesimal. When evaluating whether one untreated child is a valid counterfactual for another treated child, in the sense of satisfying **A2**, researchers should use all signal variables at their disposal. Of course, researchers never know how informative each matching variable is *ex ante*, therefore have incentives to hedge by including as much information on treatment status as possible. Doing so incurs the curse of dimensionality, and results in observations which are further and further apart as more matching variables are included. A successful application of a matching estimator needs to balance these two countervailing needs.

I use a random forest to predict the treatment status of fathers[23], and match observations using the forest's proximity matrix as my measure of how close two observations are to one another. Proximity matching has been employed in other fields to recover causal effects (e.g. Zhao et al., 2016), and fits with a larger econometric literature using random forests for causal inference and estimation of heterogenous treatment effects (Athey and Imbens, 2016; Wager and Athey, 2018; Athey et al., 2018; Friedberg et al., 2021).

The proximity matrix arises from the many decision trees which make up a given random forest. Each tree uses a randomly selected subsample of the data to categorize observations into treatment or control by sequentially partitioning the space of matching variables into hyperrectangles, known as leaves, to predict treatment status[24]. A forest then generates predictions of treatment by sorting an observation i into a leaf for each tree given its realization of observables $x_{j(i)}$, and predicting its treatment status based on whether $x_{j(i)}$ falls into leaves which contain treated or control observations the majority of the time. In this way, random forests can be understood as a class of nearest-neighbour estimators (Lin and Jeon, 2006; Biau and Devroye, 2010; Friedberg et al., 2021) and two observations i and i' with observables $x_{j(i)}$ and $x_{j(i')}$ will receive similar predictions from a forest the more often their observables fall in the same leaf.

[23] While random forests are known as the best out-of-the-box predictor in the machine learning literature, they do require researchers to manipulate tuning parameters to ensure prediction quality. In Appendix A4, I discuss my tuning procedure.

[24] For an introduction to random forests geared towards economists, see Athey and Imbens (2019). Biau and Scornet (2015) provides an excellent introduction discussing the underlying algorithms and statistical properties of random forests.

The proximity matrix summarizes how often two observations i and i' fall into the same leaves in a given random forest[25]. In particular, it is a $n \times n$ matrix which details the fraction of the trees where i and i' are sorted into the same leaf. Two observations i and i' with higher proximity will also be more similar their predicted treatment status, as a result of being close along dimensions in $\mathbf{X}$ which are predictive of treatment status. I employ this measure to uniquely match every treated observation i to the control observation i' highest in proximity, subject to a minimum proximity constraint[26].

To the best of my knowledge, proximity matching has not been used to study layoffs. Rather, the empirical literature studying layoffs has employed one of two matching procedures[27]: caliper matching (Stepner, 2019) or coarsened exact matching (Morthorst et al., 2021). In Appendix A3, I compare my matching strategy with these alternatives in detail. I employ proximity matching as it focuses on signal variables to create matches, rather than all variables made available to the procedure by the researcher. Extracting all of the information in the data on a father's future treatment status is particularly important in my setting, as my main specification in equation (4) cannot accommodate individual-level fixed effects[28]. As a result, I must rely solely on the quality of my matches to ensure my estimates satisfy **A2**.

To implement my proximity matching procedure in the data, I define the set of children who experienced parental unemployment at age a in year t as children whose father: (i) claimed UI benefits when she was age a and did not claim UI in any of the three years preceding this spell, (ii) filed taxes in all three years prior to the unemployment spell, and (iii) received positive T4 income in all three years prior to the spell. These restrictions are similar to those employed in Stepner (2019), and ensure that all treated children are

[25] Olson and Wyner (2018) explicitly relate the proximity matrix to nearest neighbour estimators by proposing the "proximity kernel" which casts random forest predictions as resulting from a kernel estimator.

[26] For most of the results presented in this paper, I restrict to a match pair with proximity greater than 0.7. This means that at least 70% of the time two observations are jointly out of bag, they appear in same leaf of a tree. While the trends in the data do not qualitatively change in response to raising or lowering this threshold by 5 percentage points, lower thresholds allow for worse matches, and run the risk of treated children being negatively selected relative to controls. This negative bias is apparent in the CATTs which arise from lower proximity thresholds. Higher thresholds further shrink my sample, and reduce my statistical power.

[27] Some past work has used mass layoffs to study parental unemployment (e.g. Oreopoulos et al., 2008). As I do not observe firms in the data, I cannot identify mass layoffs beyond a large increase in unemployment insurance claims in one Census Sub-Division.

[28] In other applications (e.g. Stepner, 2019), these fixed effects absorb time invariant heterogeneity across individuals and therefore their estimates only require balance on the trends within individuals to satisfy **A2**. As I only observe a child's income attainment once, I cannot employ these fixed effects. In Appendix A5, I present my primary test of **A2** when I match using caliper matching. When I include individual-level fixed effects, my results largely satisfy **A2**, suggesting caliper matching is sufficient for researchers whose identification relies on matching treatment and control trends within an individual. In the absence of these fixed effects, this is not the case.

experiencing an unemployment spell rather than the relatively common strategic claiming outlined by Human Resources Development Canada (1994). In particular, restriction (i) and (iii) ensure that treated fathers were steadily employed prior to the shock and did not receive UI despite being eligible to do so. Similarly, restriction (ii) ensures that fathers have complete records in the data and are not included in my estimates due to a year out of the labour market or the formal employment sector.

With this set of treated children in hand, I create set of control children who could be matched with treated children by taking all children age a in year t whose father (i) did not claim UI benefits when she was age a and nor in any of the three years preceding it, (ii) filed taxes when the child was age a and in all three years prior, and (iii) received positive T4 income when the child was age a and in all three years prior to the spell. I apply my proximity matching strategy by estimating a random forest to predict whether a given father claimed UI when his child was age a.

I follow Stepner (2019) by only including information on fathers realized in $t-3$. Using this lagged information ensures that matching variables are exogenous to treatment status, and matches are not generated using endogenous variation associated with realized treatment status. I match using information on a father's income sources[29] in year $t-3$, as well as his spouse's income sources in $t-3$ if a spouse is listed on his tax return. I also include demographic information such as province of residence, marital status, year of birth, and number of children. As my matching only explicitly restricts UI receipt in the 3 years prior to treatment, I also include a variable capturing the number of years without claiming UI. This dimension is almost always equal to the number of years between 1981 and year $t-3$.

4.2 External Validity

In Table 2, I present summary statistics on my matched sample relative to the population in year $t-3$. Columns 1 and 2 present summary statistics in $t-3$ for all fathers who claimed UI in year t, Columns 3 and 4 do the same after I remove chronic UI claimants as discussed above, and Columns 5 and 6 present summary statistics on my matched sample. Several clear trends are apparent. First and foremost, comparing Columns 1 and 3, there are no

[29]Specifically I use: total taxable income, labour income, capital gains, dividends, investment income, professional income, rental income, business income, and any CPP/QPP received. I use the same variables for spouses, as well as information on their federal and provincial taxes paid.

large differences in the income earned in year $t-3$ by all UI claimants in year t compared to those who remain in my sample after removing chronic claimants. Even though I remove 760,000 fathers who claim UI in year t to arrive at my sample of treated fathers who are possible to be matched, the only clear difference between columns 1 and 3 is in province of residence. Consistent with seasonal workers receiving UI, the chronic claimants I remove between columns 1 and 3 are more likely to live in the Atlantic provinces or Québec.

Comparing my matched sample in Columns 5 and 6 with the fathers who were possible to match in Columns 3 and 4, it is clear that my matched sample earned higher income on average than the treated observations who were possible to match. This is an artefact of the control fathers who could possibly be matched in Column 4 earning considerably higher income than treated fathers on average[30]. As a result, treated father with higher income will have many more control observations who they could possibly be matched to, and a higher likelihood of matching.

A similar force is driving fathers who are married and have higher earning spouses to be more likely to be matched. In variable importance tests, a spouse's income is highly informative in predicting a father's future unemployment in the data. Fathers whose spouses do not appear on their tax returns will lack this additional information, and will be sorted into leaves which are less predictive of their future treatment. Higher variance in the prediction results in lower proximity to control observations, relative to fathers with spouses whose future treatment can be predicted more precisely.

It is clear that my estimated treatment effects for treated children whose fathers are matched may not generalize to the children of all fathers who could have matched. Despite these external validity concerns, comparing Columns 5 and 6 suggests that my matched sample will produce internally valid results. In year $t-3$, the households treated and control children live in look almost identical in the data. In Section 4.3, I formally test whether these similarities also appear in the years before $t-3$, when household observables were not restricted by the matching procedure.

[30]In Table 2, I only provide information on net taxable income. Table A1 presents these summary statistics for total the taxable and labour income of fathers, their spouses, and their households. The same trends documented in this Section hold regardless of the income measure of interest.

4.3 Testing Identifying Assumptions

Using my matched sample, I estimate the effect of first experiencing parental unemployment at age a on a child's income attainment. As discussed in Section 2, the primary threat to identification in my setting is that unemployed fathers are negatively selected relative to their match counterfactuals, and as a result treated children are negatively selected in their potential earnings attainment.

Of course, I do not observe a child's potential earnings attainment in the data. In a standard matching design (e.g. Jager and Heining, 2019; Sarsons, 2019; Stepner, 2019), I could exploit the timing of treatment d at some time $t = \tau$ to estimate its effect an outcome y_{it}[31], and test **A2** by evaluating whether $\mathbb{E}[y_{it}|d = 0] = \mathbb{E}[y_{it}|d = 1] \;\; \forall t < \tau$ in my matched sample. As $y_{it} = y_{it}(0)$ for all periods $t < \tau$ in these settings, it is reasonable to argue that if conditional independence holds for periods before τ then it ought to for $t \geq \tau$ as well.

Unfortunately, in the case of income attainment I only observe the child's outcome once, and over a decade after most children were treated. Instead of testing differences in the child's outcomes prior to treatment, I test for differences between the *households* treated and control children are exposed to prior to treatment. My matching strategy places restrictions on treated and control fathers in the 3 years prior to treatment. As I do not place any restrictions on the data in periods before the child was age $a - 3$, in year $\tau - 3$, and I implement this test for all children treated at age a and their controls by estimating a regression of the form,

$$\theta_{j(i)t} = \sum_t \mu_t + \delta_t D_i(a) + \epsilon_{j(i)t} \tag{5}$$

where $D_i(a) = 1$ if a child is treated at age a and $D_i(a) = 0$ for a matched counterfactual who is not treated at a. As I estimate separate treatment effects for each age, I run regressions with the same form as equation (5) separately for my matched sample using children treated at each age $a \in \{2, 3, ..., 21\}$ and their matched counterfactuals.

I evaluate whether treated and control households differed prior to treatment by jointly testing $H_0 : \delta_t = 0 \;\; \forall t < \tau - 3$. To illustrate the test, Panel A of Figure 1 plots the raw pretrends in father's log taxable income for children treated age 6, while Panel B presents

[31] Assume that $y_{it} = y_{it}(0)(1 - d\,\mathbf{1}\{t \geq \tau\}) + y_{it}(1)(d\,\mathbf{1}\{t \geq \tau\})$

the resulting estimates of δ_t. As the majority of children in my sample were born before 1981, I implement this test up to a minimum age $\underline{a} = 1981 - \text{YOB}$ for each year of birth (YOB). For the cohorts born after 1981 I consider all parental observables occurring from 1981 onwards, despite the implied $\underline{a} < 0$, as a test for possible differences in fertility decisions in response to income dynamics between treatment and control.

In Table 4, I present the p-values which result from testing $H_0 : \delta_t = 0 \quad \forall t < \tau - 3$ for a number of household observables $\theta_{j(i)t}$. I test for differences in a father's or a father's spouse's total taxable, labour, and net of tax income, and reject eight null hypotheses of no difference in earnings prior to $\tau - 3$ at the 5% level. Looking at whether the father was married or moved to a different Census Sub-Division, I reject two additional null hypotheses at the 5% level. Finally, focussing on differences in the tax-filing behaviour for fathers, their spouses, and whether the father's spouse was in the labour market, I reject one of the null hypotheses at the 5% level. Of the 220 hypothesis tests performed in Table 4, I reject the null hypothesis of no observable differences between treated and control households prior to treated 11 times at the 5% level, in line with expected Type I error rates[32].

In light of these results, it appears that there is little reason to believe to that control households differed from treated households prior to treatment status being realized in year τ at with their children aged a. One concern with estimates of δ_t in equation (5) is that the variation between households may be swamping differential trends within households. To ensure this is not the case, and for consistency with similar work employing matching estimators in the literature, I present an alternative version of Table 4 in Appendix Table A4 which allows for parent-level fixed effects in equation (5)[33]. I reject 9 of the 220 tests at the 5% level, again well in line with expected Type I error rates.

5 Results

In this Section, I present my estimates of the effect of first experiencing parental unemployment at age a on a child's income attainment, Y_i. I also present estimates for the effect of parental unemployment at experienced age a on alternative outcomes such as whether the

[32] I reject none of these hypotheses at the 1% level and 17 at the 10% level, again in line with the expected Type I error rate.

[33] In particular, I run $\theta_{j(i)t} = \chi_{j(i)} + \sum_{t \neq \tau-3} \tilde{\mu}_t + \tilde{\delta}_t D_i(a) + \epsilon_{j(i)t}$ where $\chi_{j(i)}$ is a parent-level fixed effect, and test $H_0 : \tilde{\delta}_t = 0 \quad \forall t < \tau - 3$. Following the literature, all p-values presented result from clustering at the child level (Jager and Heining, 2019).

child attended a post-secondary educational institution, took up more income supplements from the social safety net, or had children earlier. I estimate these treatment effects using the regression specification presented in equation (4), re-stated here for convenience,

$$Y_i = \lambda_a + \Delta_a D_i(a) + x_i'\gamma_a + \epsilon_i$$

where x_i are child-level controls which were not restricted in the matching procedure and are independent of treatment, but may affect earnings, such as the child's sex. In my preferred specifications, the only control I include is the child's sex and I restrict $\gamma_a = \gamma \; \forall a$ to maximize statistical power, though the results are qualitatively robust to alternative specifications.

These reductions in Y_i, $\hat{\Delta}_a$, are caused by first experiencing parental unemployment at age a and correspond directly to my estimands from Section 2 in equations (2) and (3). As such, my estimated effects are best understood as CATTs: conditional on first experiencing parental job loss at age a, they are the average reduction in Y_i for treated children. Each estimated effect averages over a number of treatment effects which depend on the treated father's potential outcomes for employment in the future, as defined in Section 2.1. When averaging, each of these treatment effects are weighted by the likelihood of observing a particular sequence $(U_\alpha(0), U_\alpha(1))_{\alpha=a,a+1,\dots,21}$ for a treated father.

To maximize statistical power statistical power in equation (4), I consider alternative specifications which bin multiple ages of treatment together. As I observe treatment for children exposed at 20 separate ages, one natural approach is to use 10 equally sized bins of children treated at two different ages (i.e. binning 2 and 3 together, 4 and 5, etc.) rather that a for a single age. For robustness, I also consider bins of three different ages of exposure, creating 6 equal bins (4 to 6, 7 to 9, etc.) and a single residual bin for children treated at ages 2 and 3.

5.1 Income Attainment

Figure 2 presents my estimates of Δ_a along with 95% confidence intervals by age of exposure, binned into three year age bins for statistical power. The exact figures are provided in Table 3, along with standard errors clustered within a match pair as suggested by Abadie et al. (2017). In Figure A2 and Tables A2 and A3, I also present estimates of Δ_a by each age of exposure to paternal unemployment and two year age bins. I consider both income attainment at ages 27 to 31 and ages 30 to 34 to ensure that the lifecycle bias is not affecting the results, however consistent with Chetty et al. (2014) there appears to be very little qualitative or quantitative difference in the results from either outcome.

Consistent with when children are most dependant on their parents, Figure 2 shows that children who experienced parental unemployment at any point before their late teens tended to attain 1 to 4 rank points lower than the relevant counterfactuals. Focussing on income attainment at ages 27 to 31, there is a clear trend among the estimated effects. Children treated at the youngest ages experienced large and significant reductions in income attainment, around 3 to 4 rank points. Slightly older children treated around the ages of 5 or 6 show little evidence of lower income attainment than the control group, experiencing a reduction of about 1 rank point which is generally not statistically significant at conventional levels. Children treated at ages 7 or 8 tend to experience a 3 rank point reduction in earnings attainment, while children exposed to parental unemployment around ages 15 to 16 experience a 2 rank point reduction. Consistent with when chidren begin to leave the household, I find little evidence that parental unemployment affects a child's income attainment for children exposed to parental unemployment after age 18.

In Figure 3, I fit a quartic polynomial of age on my treatment effects. For reference, I also plot my treatment effects using three year age bins. The quartic fit presents the same trend as my non-parametric estimates of the effect of parental unemployment on a child's income attainment. The treatment effects are weakly increasing the age of exposure for children treated before age 5, and then weakly decreasing from ages 5 to 10. While there is little evidence that children treated between the ages of 2 and 10 experience different losses in income attainment, there is considerable evidence that children exposed after age 11 begin to experience smaller and smaller losses in income attainment. For children treated in their late teens or early twenties there is little evidence of any reduction in their income attainment, consistent with their parents beginning to have less of an effect on their human capital accumulation.

I consider several methods to provide statistical inference on this trend in treatment effects. I cannot reject the null hypothesis that experiencing parental unemployment in neighbouring ages bins has the same effect on a child's income attainment (i.e. $H_0 : \Delta_{7-9} \geq$

Δ_{10-12}). To maximize statistical power, I follow work on neighhourhoods (Chetty et al., 2016) and test whether being first treated at or before age 12 leads to a larger reduction in income attainment than being treated after age 12 (i.e. $H_0 : \Delta_{2-12} \geq \Delta_{13-21}$). This test yields a p-value of 0.0727, suggesting that children treated at younger do indeed experience larger reductions in earnings attainment, albeit only at the 10% level. Looking at Panel A of Figure 3, there appears to be a large difference between the effects of parental unemployment on children exposed before and after age 10. These are also the ages which the literature had previously been unable to study. Testing whether these children exposed to parental unemployment before age 11 experience larger reductions in income than children treated at older ages, $H_0 : \Delta_{2-10} \geq \Delta_{11-21}$, I can reject the null at the 5% level (p-value 0.0472).

In Panel (b) of Figure 3, I test whether there is a monotonic and increasing relationship between the age of exposure to parental unemployment and the reduction in income attainment it causes. I present a nonparametric, local linear regression specification of the trends in my treatment effects by age, constrained to be weakly increasing by rearrangement following Chernozhukov et al. (2009). Alongside the estimated relationship between age of exposure and income attainment losses, I plot a bootstrapped 95% uniform confidence interval obtained from 1,000 replications, with re-sampling stratified at the match-level following Abadie et al. (2017). In this Panel, I can test the null hypothesis that there is no relationship between age of exposure to parental unemployment and the resulting reduction in the child's income attainment. In particular, if I am unable to plot a horizontal line through this interval, then I can reject the null hypothesis of no monotonicity at the 5% level. Panel (b) of Figure 3 presents one such horizontal line which can be plotted through the confidence interval, as such I fail to reject the null hypothesis that the relationship between age of exposure to parental unemployment and a child's income attainment is monotone.

Taken together, Panels (a) and (b) of Figure 3 provide evidence that children exposed to parental unemployment at age 10 or younger experience larger reductions in their income attainment than children exposed at older ages. Despite these average differences, I fail to reject the null hypothesis that there is not a monotonic relationship between age of exposure and reduced earnings attainment. Figure 2 is informative in reconciling these two results. Children exposed to parental unemployment around ages 5 to 6 experience considerably smaller reductions in earnings attainment than children exposed at neighbouring ages. As

Panel (a) of Figure 3 uses a quartic polynomial of age, it is able to capture this local maximum in my estimated effects. Despite this non-monotonicity around the age children enter school, experiencing parental unemployment at younger ages still leads to larger reductions in income attainment overall.

Results are similar for income attainment at ages 30 to 34, although estimates are only available for children treated at ages 5 to 21 due to data constraints. Again, there is little statistical evidence of effect of parental unemployment on children treated at ages 5 or 6, though these estimates lack the power to detect any estimates in the range of a 1 to 2 rank point reduction in earnings. Children treated at ages 7 to 10 experience large losses in income attainment due to their father's unemployment, in the range of 3 to 4 rank points. For children treated at age 10 onwards, the estimates gradually shrink from a 2 rank point reduction in earnings through to a null effect around ages 18 to 21.

To the best of my knowledge, the literature had previously not documented the effects of parent unemployment on a child's income attainment for children treated before age 10. As a result, this is is the first evidence of a plateau in treatment effects for children between the ages of 2 and 12 whose parents lost their jobs. Similarly, the literature had limited systematic evidence of smaller and smaller treatment effects for children treated between the ages of 13 and 21.

Both of these results have considerable policy implications. First, in the Canadian context, modern Employment Insurance pays Supplementary Benefits to parents with children under age 6. My results suggest that children treated between the ages of 7 and 12 experience roughly the same reduction in income attainment as children treated before 6. As such, Canadian policymakers should carefully consider extending Supplementary Benefits to parents with dependants older than 6 as well.

Second, for jurisdictions where employment insurance does not vary with the ages of dependants, I document systematic differences in the effects of parental unemployment on children by their ages of exposure. While I do not explicitly model search effort in this paper, these dynamics suggest that parents with younger dependants suffer larger losses from unemployment than parents with older dependants. If that is the case, as such parents with younger dependants will search for work with effort which is more inelastic to unemployment insurance benefits than those with older dependants and the optimal benefit schedule ought to take this into account[34].

These trends fit with a general timeline of childhood and education in Canada. Before age 4, child care must be privately provided by parents in some form[35] and these ages are often understood to be critical periods in the child's development. At ages 4 to 5, children are able to attend kindergarten in public schools and enter Grade 1 at age 6. Children are able to attend public school until age 17 or 18 depending on the province of residence, when they enter the labour force or attend a post-secondary institution at some private costs. If paternal unemployment affects the household's ability to provide child care, we might expect that large effects manifest themselves at younger ages when the household provides all care privately. Once the household receives publicly provided childcare for part of the work day, the effects of unemployment on child may be reduced. Similarly, if the skills children learn in elementary school like reading comprehension provide a base for their future human capital accumulation in any sector, then parental unemployment may have larger impacts if it affects children at the age when they learn these skills.

Overall, my results suggest parental unemployment causes loss of 3 to 4 percentile ranks in income attainment for children treated before age 12. As discussed in the introduction, losses on this scale mean that parental unemployment has a considerable effect in the long-run on children exposed at young ages. In Canada, Connolly et al. (2019) estimate a rank-rank correlation between a parent and a child's income of 0.21. If I take their estimate as the causal effect of parental income on children, then the 3.4 percentile rank reduction in income attainment at ages 27 to 31 for children treated ages 2 to 3 is equivalent to these children growing up in a household which was 16.2 percentiles lower in the income distribution. Similarly, Stepner (2019) estimates that layoffs cause an average reduction in after tax income of 10% (approx. $8,000) over five years. I find that experiencing parental unemployment at ages 6 or 7 causes a 6% ($2,500) reduction in average earnings attainment between the ages of 30 and 34. In this case, parental unemployment experienced early in childhood causes a reduction in earnings attainment roughly two thirds the size of the income losses caused by the child *themselves* being laid off.

[34]Note that assuming search does not vary with the child's age is inconsistent with forward-looking parents. Indeed, if parents are aware of the dynamic spillovers which result from unemployment in a single period, then their search behaviour should vary systematically with the age of their child. In other work (Uguccioni, 2020), I consider how these dynamic spillovers affect search decisions and optimal unemployment insurance, as they not consistent with Assumption 5(c) in Chetty (2006).

[35]In 1997, Québec began to introduce Canada's first universal daycare at a highly subsidized rate of $ 5 per day (Baker et al., 2008), however my all of the children in my sample were too old to attend this care by then.

5.2 Socio-Economic Outcomes

It is clear that parental unemployment has large, long-run effects on a child's income attainment. To understand more broadly how parental unemployment affects a child's well-being in the long-run, it is important to consider additional socio-economic outcomes. To ensure comparability of the estimates, I use the same regression specification as equation (4), substituting in different socio-economic variables as my outcome of interest rather than income attainment.

One natural socio-economic indicator to consider is whether parental unemployment affects a child's educational attainment. While the literature has documented that untimely parental unemployment can affect high school completion (Tanndal and Paallysaho, 2020), there is some debate whether it affects a child's likelihood of attending post-secondary (Hilger, 2016). I find that parental unemployment does cause children exposed at young ages to be less likely to attend post-secondary education.

As discussed in Section 3.1, I only observe post-secondary tuition tax credits after 1997 in the data, and as a result I only observe younger birth cohorts at their prime post-secondary ages. In Figure 4, I present estimates of Δ_a when using post-secondary attendance as my outcome of interest. Treated children are 3 to 4 percentage points less likely to attend post-secondary than the control, compared to a mean attendance rate of 80% in the control group. Of course, I do not observe whether a child completes her post-secondary degree in the data. If we expect that there is a monotone relationship between the effects of parental unemployment on post-secondary attendance and completion, however, then my results are suggestive of parental unemployment affecting the educational attainment of children[36].

Another important indicator of socio-economic status is the receipt of social safety net benefits. Indeed, receiving UI and social assistance benefits respectively indicate a need for support against labour market uncertainty and poverty. Beyond indicating financial need, receipt of social insurance benefits such as UI has also been documented to carry a social stigma (Brand, 2015). In Figures 5 and 6, I consider whether a child ever receives social assistance and UI benefits in adulthood.

[36]Alternatively, these differences may occur even if there is no difference between a treated and control child's post-secondary attendance. For example, treated children may be less likely to be aware of these tax credits and less likely to claim them as a result. Indeed, some issues with the salience of these credits has been documented Frenette (2017). If we assume that there is no systematic difference in tax credit salience depending on treatment status, then my estimates will exclusively reflect a difference in post-secondary attendance. While this is not presently testable in the IID, it will be testable in the forthcoming linkage between the IID and Census data (Connolly et al., 2021) where educational attainment is observed.

In Figure 5, I consider whether a child ever received social assistance between the ages of 25 and 31, and 30 and 34 as my outcomes of interest. Children treated before age 15 were approximately 1.9 percentage points more likely to receive social assistance than the control group between the ages of 25 and 31, compared to a mean receipt rate of 5.1% in the control group. This gap narrows somewhat at older ages, as children treated before age 15 are only 1.3 percentage points more likely to receive social assistance between the ages of 30 and 34, compared to a mean receipt rate of 4.0% in the control group. Especially given the low social assistance use in the control group, these results suggest that parental unemployment causes treated children to be considerably more likely to face financial insecurity and poverty.

In Figure 6, I consider whether a child ever received UI benefits between the ages of 25 and 31, and 30 and 34 as my outcomes of interest. As with social assistance, treated children are considerable more likely to receive UI benefits than the control group between the ages of 25 and 31. Again, the gap narrows when looking at UI receipt between the ages of 30 and 34. Unlike with social assistance receipt or income attainment, I find that treated children are more likely to receive UI benefits between the ages of 25 and 31 *regardless* of their age of exposure. Similar to my results for social assistance, these results suggest parental unemployment causes children to be much more likely to face financial insecurity. As treated children are more likely to receive unemployment insurance benefits, one source of this financial insecurity appears to be treated children finding jobs in much more unstable work environments[37].

My results are in line with previous work in Canada using older birth cohorts. Focussing on children who experience parental unemploymeny between the ages of 10 and 14, Oreopoulos et al. (2008) also find that treated children are 1.5 percentage points more likely to receive social assistance than their control group and 3.9 percentage points more likely to receive unemployment insurance benefits between the ages of 25 and 29.

In addition to its effects on their income attainment, parental unemployment causes children to be less likely to attend post-secondary schooling, more likely to receive social assistance benefits, and more likely to receive UI benefits. In Table 5, I consider whether treated children go on to live in households with different structures than the control children. As outcomes, I consider whether a child was married at ages 25 and 30, had any children by

[37] Of course, modern Canadian EI still pays benefits to seasonal workers. Whether these children are more likely to be employed in seasonal jobs or jobs at a higher risk of unemployment remains to be seen.

age 30, and their age at the first birth of their own child. There is little evidence of any differences between treated and control children along any of these dimensions, regardless of their age of exposure to parental unemployment.

6 Mechanisms

Parental unemployment is a complex shock for children to face, potentially affecting their futures through channels other than their household's lost resources. Understanding whether channels such as familial stress and moving to worse neighbourhoods play a role in how parental unemployment affects children is crucial for developing policy prescriptions to mitigate these effects.

In this Section, I investigate two related mechanisms to understand how parental unemployment affects children. First, I explore whether the timing of household income losses explains why children treated at younger ages tend to experience larger reductions in income attainment. Second, I consider how much of the overall treatment effects can be explained by parental unemployment causing lost household income, changes in neighbourhood quality (proxied with fraction of adults holding a bachelor's degree), and household stress (proxied by parental divorce). For both of these exercises, I employ a mediation analysis as proposed by Heckman et al. (2013).

In Appendix A6, I consider whether my application satisfies the additional assumptions in Heckman et al. (2013) required for my decomposition to have a causal interpretation. I argue that my results are best understood as suggestive evidence, rather than causal results. In this exercise, I consider how the losses in household inputs to the child's human capital which result from parental unemployment, such as neighbourhood quality, relate to a child's income attainment. To document the lost inputs, $\bar{\theta}_{j(i)}$, I run regressions of the form,

$$\bar{\theta}_{j(i)} = \sum_{a} \phi_a D_i(a) + x_i'\Gamma + \mu_{ii'} + \varepsilon_i \tag{6}$$

where $\hat{\phi}_a$ is average loss in input $\bar{\theta}_{j(i)}$ experienced by children who first experience parental unemployment at age a. I then estimate the average return on these inputs by age of exposure, ψ_a,

$$Y_i = \sum_a \tau_a D_i(a) + \psi_a \bar{\theta}_{j(i)} + x_i'\gamma + \mu_{ii'} + e_i \tag{7}$$

Ultimately, Frisch-Waugh-Lovell theorem implies that I can decompose my treated effects into a portion explained by lost inputs $\bar{\theta}_{j(i)}$ and a residual, explained by losses in unobserved inputs to the child's income attainment. Specifically,

$$\Delta_a = \underbrace{\tau_a}_{\text{Residual}} + \underbrace{\psi_a \phi_a}_{\text{Explained}}$$

It is important to notice that I estimate these differences with match pair fixed effects, $\mu_{ii'}$. As all matches in my sample are one to one, these fixed effects can be dropped from equation (6), or included in equation (4), without changing my estimates of $\hat{\phi}_a$ or $\hat{\Delta}_a$. As discussed in Appendix A6, these fixed effects are important for identification of $\hat{\psi}_a$ in equation (7), as they ensure that cross-sectional differences prior to treatment do not affect my decomposition.

6.1 Timing of and Exposure to Household Income Losses

To investigate how the timing of parental unemployment affects children, I consider how the father's income scarring relates to the child's income attainment in adulthood. In the context of equation (7), this amounts to estimating ψ_a with different income concepts.

One approach to investigate whether the timing of unemployment drives results is to fix the number of years of earnings and consider how the return on these earnings varies by the timing of treatment. Figure 7 presents two income variables in this vein. In both panels of the Figure, the number of years of earnings used as an explanatory variable is fixed so children treated at younger ages do not experience mechanically larger losses as a result of experiencing more years of scarred earnings. Panel A presents estimates of ψ_a resulting from regressing the father's income rank in the year of unemployment and the year immediately thereafter on his child's income attainment at ages 30 to 34, conditional on match pair fixed effects and the child's sex. Panel B does the same, employing the father's income rank at ages 16 to 18 in the regressions instead of his earnings rank immediately after the unemployment spell.

The point estimates presented in Panel A of Figure 7 are best interpreted as the relationship between the immediate income losses from parental unemployment and his child's income attainment. For the children treated at 7 to 9 years old in Panel A, for example, a 10 percentile drop in father's earnings immediately after unemployment relative to the counterfactual is associated with a 1.54 percentile decline in the child's earnings. For children treated at younger (older) ages, the point estimates reflect income losses experienced at younger (older) ages. As Panel A presents point estimates which are generally decreasing with age of treatment, it appears that income losses of a given size experienced at younger ages tend to be associated with larger reductions in the child's income attainment. This trend is consistent with critical periods in human capital investments which tend to occur at younger ages, resulting in higher returns on income and larger losses from parental unemployment at these ages.

Unlike Panel A of Figure 7, Panel B presents the relationship between paternal earning at a fixed age and the child's income attainment in adulthood. In this case, children treated at young ages will be living in households dealing with long-run income scarring when they are ages 16 to 18, while children treated just before or at ages 16 to 18 will be experiencing the immediate effects of parental unemployment. Again there is a general trend in these figures of higher returns on parental income for children treated at younger ages, however these may reflect dynamic complementarities in human capital investments rather than critical periods of investment. In particular, if human capital investments are complements then forgoing investments at younger ages reduces the returns on future investments. These results are suggestive of large complementarities in the foundational skills formed at young ages and the return on parental income at ages 16 to 18.

The results in Figure 7 are suggestive of critical periods in human capital investments made at younger ages and important dynamic complementarities between the investments made at younger ages and future investments in children. They should not, however, be interpreted as causal estimates of ψ_a from equation (7). Income losses in the first 2 years following an unemployment spell are correlated with the long-run income losses which follow, and may confound critical periods with overall losses. Similarly, the income scarring experienced ages 16 to 18 for children treated at young ages is correlated with the size of the initial income loss, and may contaminate results.

Figure 8 remedies these shortcomings using the father's log average earnings in all years between the unemployment spell and the child reaching age 21 rather than income earned in a fixed window. Moreover, using a log income variable ensures that the paternal income comparisons within a treat-control match pair are proportionate, unlike the rank-rank specifications where a 10 percentile point reduction in the father's earnings differs in scale and proportionate income loss depending point of reference. If the overall exposure to total parental income loss determined the child's income attainment, we should expect that the relationship between log average earnings at the child's income attainment did not have a varying slope across all ages of exposure. In other words, every 1% increase in the father's average earnings should pay a relatively constant return in the average year of exposure, rather than paying higher returns at younger ages.

If the rate of return varies by age of exposure, then either critical periods of investments or dynamic complementarities are at play. For example, consider comparing the average return on parental income faced by two children treated at ages 5 and 12 relative to their respective counterfactuals in all years after they are treated. Both children will experience reduced income relative to their controls at ages 10 to 21, which could yield the same return if there are no dynamic complementarities in investment. If there are dynamic complementarities in investment, then the reduced income experienced by the younger treated child ages 5 to 11 will mean that investments made at ages 12 to 21 should pay a *lower* rate of return than the older child who did not experience reduced parental earnings at ages 5 to 11. In that case, we should expect that the average return on parental income for children treated at older ages is higher for children treated at older ages. Alternatively, if the rate of return on paternal earnings is decreasing with age, then it is likely the case that there are critical periods in human capital investments, and investments at younger ages pay higher returns than those made at younger ages.

In Figure 8, the relationship between father's log average income after unemployment and the income attainment of his child is decreasing with age of exposure. This trend is inconsistent with dynamic complementarities in human capital investments or overall exposure being the primary driver of the decreasing Δ_a estimates presented in Figure 2. These differences in returns can explain large differences in the treatment effects Δ_a. The average child treated at age 5 had a father who earned approximately 21.3% less then his relevant counterfactual. For every 10% reduction in earnings within a match pair, children who were treated at ages 7 to 9 attained an income rank 0.51 percentiles lower then her counterfactual whereas children treated at ages 16 to 18 attained a rank only 0.22 percentiles lower. At the same proportionate loss of 27.4% of paternal income, these estimates imply that children treated at ages 7 to 9 will lose 0.78 percentile ranks more than children treated at age 16 to 18.

6.2 Mediation Analysis

Using the Heckman et al. (2013) decomposition, I can express how much of the effect of parental unemployment on a child's income attainment is attributable to losses in a given input, ϕ_a, by multiplying the losses with the estimated return on that input, ψ_a. I consider three inputs to a child's income attainment which have been documented in the literature. To capture how lost household resources affect children, I use father's log average income after unemployment. I use the average fraction of adults holding a bachelor's degree in the Census Sub-Division where the child resides after experiencing parental unemployment to proxy for neighbourhood quality. Finally, I capture stress on the household using the incidence of parental divorce following parental unemployment.

In Table 6, I present my mediation analysis. For each input to the child's income attainment, I report (i) the change in the input age parental unemployment relative to the control group, (ii) the estimated return on the input, and (iii) the percentage of my treatment effect explained by the that particular input. Looking at losses in a father's income for children treated between the ages of 7 and 9, for example, Column (1) shows that treated fathers earn roughly 27.4% less than control fathers in total income between the year they are employed and the year their child turns 21. Column (2) reports the return on these reduced earnings, conditional on neighbourhood quality and the incidence of divorce. For every 10% reduction in the fathers earnings, children experience a loss of 0.499 percentile ranks in their birth cohort's income distribution. Altogether, the 27.4% reduction in the father's earnings times a return of 4.99 accounts for a 1.37 percentile reduction in the child's income rank in adulthood — 41% of my estimate effect for these children, as reported in column (3).

Overall, reductions in the father's income tend to account for roughly 40% of my estimated treatment effects. Reductions in neighbourhood quality an additional 5% of effects, while an increased incidence of parental divorce explains little to none of the effects of parental unemployment on children. While these figures cannot be interpreted as causal, they are suggestive of which policies could mitigate the long-run effects of parental unemployment on children.

Given the declining return on parental income by age of exposure, my results suggest that policy makers should consider paying higher unemployment insurance benefits to parents with younger children. Similarly, as income scarring explains the largest fraction of my results, my results suggest that policy makers may wish to target re-training policies at young parents (e.g. by providing child-care) whose children stand to benefit the most.

Focussing on avenues for targeted relief, Table 6 presents evidence that parental unemployment causes families to move to less educated neighbourhoods, and that these moves explain around 4% of my treatment effects. In light of the extensive work studying the importance of neighbourhoods in a child's development (Chetty and Hendren, 2018a; Chyn, 2018; Laliberté, 2021), one policy goal may be to help parents avoid uprooting their families in response to unemployment[38]. In that case, mandatory mortgage deferrals in the event of a layoff, such as those proposed by the Liberal Party during the 2021 Canadian federal election (Liberal Party of Canada, 2021), could also help alleviate the intergenerational effects of unemployment.

Finally, parental divorce has been documented to have no affect on children in the long-run (Corak, 2001), and my results should not suggest that household stress resulting from parental unemployment does not affect children in the long-run. As other work has shown that the stress resulting from parental unemployment affects a child's educational attainment (Tanndal and Paallysaho, 2020), it is likely the case that the incidence of divorce is not a good proxy for household stress. Future work should strive to make use of data with better information on the mental health of children and their parents to study whether it explains more of the long-run effects of parental unemployment on children.

7 Conclusion

Parental unemployment can have large, long-run effects on children. Using a random forest matching procedure, I estimate the causal effects of parental unemployment on a child's

[38] Of course, if low opportunity areas also have more job vacancies, then introducing policies which discourage mobility may not be optimal for the planner. This is an area which requires further exploration in the literature.

income attainment by age of exposure. I provide the first empirical evidence of the long-run effects of parental unemployment on children exposed between the ages of 2 and 10, a period thought to be critical for child development. I find that children exposed to parental unemployment at young ages experience losses of 3 to 4 rank points in average earnings attainment in adulthood (approximately $2,500 per year). These children are also 36% more likely to receive welfare as adults and 4% less likely to pursue post-secondary education. Consistent with critical periods of child development, children who experience parental job loss before age 10 experience larger reductions in income attainment than children exposed at older ages.

Decomposing my estimates, I find that parental income scarring accounts for 30% to 50% of my estimated treatment effects, while moves to lower oppportunity neighbourhoods explain an additional 4% of these effects. My results are suggest that policy makers should consider whether short term relief such as unemployment insurance benefits which vary with the ages of dependants or mandatory mortgage deferrals could help mitigate these effects. Future work should strive to further investigate how the dynamics in a parents employment status and income scarring contribute to the effects of parental unemployment on children. In Section 2, I show that my overall estimates can be understood as weighted averages of a number of treatment effects which depend on whether parents exhibit state dependence in their employment status. Identifying these weights using methods such as those proposed by Torgovitsky (2019) will help inform whether duration dependence plays an important role in my results, better informing which policy levers are available to mitigate the intergenerational effects of unemployment.

References

Steven J Davis and Till Von Wachter. Recessions and the costs of job loss. *Brookings Papers on Economic Activity*, (2):1–72, 2011. doi: doi:10.1353/eca.2011.0016.

Michael Stepner. The insurance value of redistributive taxes. Working paper, University of Toronto, 2019. URL http://files.michaelstepner.com/Stepner_JMP_2019-11-10.pdf.

Sarah A. Burgard, Jennie E. Brand, and James S. House. Toward a better estimation of the effect of job loss on health. *Journal of Health and Social Behavior*, 48(4):701–721, 2007. doi: doi:10.1177/002214650704800403. URL https://pubmed.ncbi.nlm.nih.gov/18198685/.

Jennie E. Brand and Sarah E. Burgard. Job displacement and social participation over the lifecourse: Findings for a cohort of joiners. *Social Forces*, 87(1):211–242, 2008. doi: 10.1353/sof.0.0083. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2935181/.

Jennie E. Brand. The far-reaching impact of job loss and unemployment. *Annual Review of Sociology*, 41(1):359–375, 2015. doi: 10.1146/annurev-soc-071913-043237. URL https://doi.org/10.1146/annurev-soc-071913-043237.

Marius Opstrup Morthorst, David J. Price, and Peter Rønø Thingholm. The effects of layoffs on opioid use and abuse. Working paper, University of Toronto, 2021.

Kerwin Kofi Charles and Melvin Stephens. Job displacement, disability, and divorce. *Journal of Labor Economics*, 22(2):489–522, 2004. ISSN 0734306X, 15375307. URL http://www.jstor.org/stable/10.1086/381258.

Ariel Kalil and Kathleen M. Ziol-Guest. Single mothers' employment dynamics and adolescent well-being. *Child Development*, 76(1):196–211, 2005. ISSN 00093920, 14678624. URL http://www.jstor.org/stable/3696718.

Nathaniel G. Hilger. Parental job loss and children's long-term outcomes: Evidence from 7 million fathers' layoffs. *American Economic Journal: Applied Economics*, 8(3):247–83, July 2016. doi: 10.1257/app.20150295. URL https://www.aeaweb.org/articles?id=10.1257/app.20150295.

Julia Tanndal and Miika Paallysaho. Family-level stress and children's educational choice: Evidence from parental layoffs. Working paper, Brown University, 2020. URL http://juliatanndal.com/JMP.pdf.

Jenifer Ruiz-Valenzuela. The effects of parental job loss on children's outcomes, 08 2021. URL https://oxfordre.com/economics/view/10.1093/acrefore/9780190625979.001.0001/acrefore-9780190625979-e-654.

Philip Oreopoulos, Marianne Page, and Ann Huff Stevens. The intergenerational effects of worker displacement. *Journal of Labor Economics*, 26(3):455–000, 2008. doi: 10.1086/588493. URL https://doi.org/10.1086/588493.

Eva Mork, Anna Sjogren, and Helena Svaleryd. Consequences of parental job loss on the family environment and on human capital formation: Evidence from plant closures. Working Paper 12559, IZA - Institute of Labour Economics, August 2019. URL https://www.iza.org/publications/dp/12559/consequences-of-parental-job-loss-on-the-family-environment-and-on-human-capital-formation-evidence-

Raj Chetty, Nathaniel Hendren, and Lawrence F. Katz. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106 (4):855–902, April 2016. doi: 10.1257/aer.20150572. URL https://www.aeaweb.org/articles?id=10.1257/aer.20150572.

Raj Chetty and Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects*. *The Quarterly Journal of Economics*, 133(3):1107–1162, 02 2018a. ISSN 0033-5533. doi: 10.1093/qje/qjy007. URL https://doi.org/10.1093/qje/qjy007.

Raj Chetty and Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates*. *The Quarterly Journal of Economics*, 133(3):1163–1228, 02 2018b. ISSN 0033-5533. doi: 10.1093/qje/qjy006. URL https://doi.org/10.1093/qje/qjy006.

Eric Chyn. Moved to opportunity: The long-run effects of public housing demolition on children. *American Economic Review*, 108(10):3028–56, October 2018. doi: 10.1257/aer.20161352. URL https://www.aeaweb.org/articles?id=10.1257/aer.20161352.

Janet Currie and Douglas Almond. Chapter 15 - human capital development before age five. volume 4 of *Handbook of Labor Economics*, pages 1315–1486. Elsevier, 2011. doi: https://doi.org/10.1016/S0169-7218(11)02413-0. URL https://www.sciencedirect.com/science/article/pii/S0169721811024130.

Douglas Almond, Janet Currie, and Valentina Duque. Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446, December 2018. doi: 10.1257/jel.20171164. URL https://www.aeaweb.org/articles?id=10.1257/jel.20171164.

James Heckman, Rodrigo Pinto, and Peter Savelyev. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–86, October 2013. doi: 10.1257/aer.103.6.2052. URL https://www.aeaweb.org/articles?id=10.1257/aer.103.6.2052.

Miles Corak. Death and divorce: The long-term consequences of parental loss on adolescents. *Journal of Labor Economics*, 19(3):682–715, 2001. doi: 10.1086/322078. URL https://doi.org/10.1086/322078.

Marie Connolly, Miles Corak, and Catherine Haeck. Intergenerational mobility between and within canada and the united states. *Journal of Labor Economics*, 37(S2):S595–S641, 2019. doi: 10.1086/703465. URL https://doi.org/10.1086/703465.

Bruce Sacerdote. How large are the effects from changes in family environment? a study of korean american adoptees. *The Quarterly Journal of Economics*, 122(1):119–157, 2007. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/25098839.

Andreas Fagereng, Magne Mogstad, and Marte RÃžnning. Why do wealthy parents have wealthy children? *Journal of Political Economy*, 129(3):703–756, 2021. doi: 10.1086/712446. URL https://doi.org/10.1086/712446.

Martti Kaila, Emily Nix, and Krista Riukula. Disparate inpacts of job loss by parental income and implications for intergenerational mobility. Working paper, University of Southern California, 2021. URL https://drive.google.com/file/d/1nLYV3gf7jOgA8xjzZdJE4PIwJaNeyGJn/view.

Orla Doyle. The first 2,000 days and child skills. *Journal of Political Economy*, 128(6):2067–2122, 2020. doi: 10.1086/705707. URL https://doi.org/10.1086/705707.

Hilary Hoynes, Diane Whitmore Schanzenbach, and Douglas Almond. Long-run impacts of childhood access to the safety net. *The American Economic Review*, 106(4):903–934, 2016. ISSN 00028282. URL http://www.jstor.org/stable/43821480.

Anna Aizer, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney. The long-run impact of cash transfers to poor families. *American Economic Review*, 106(4):935–71, April 2016. doi: 10.1257/aer.20140529. URL https://www.aeaweb.org/articles?id=10.1257/aer.20140529.

G. S. Becker and Nigel Tomes. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87(6):1153–1189, 12 1979.

Gary S. Becker and Nigel Tomes. Human capital and the rise and fall of families. *Journal of Labor Economics*, 4(3):1, Jul 01 1986. Last updated - 2013-02-23.

Gary Solon. Intergenerational mobility in the labor market. volume 3 of *Handbook of the Labor Economics*, pages 1761 – 1800. Elsevier, 1999.

Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *. *The Quarterly Journal of Economics*, 129(4):1553–1623, 09 2014. ISSN 0033-5533. doi: 10.1093/qje/qju022. URL https://doi.org/10.1093/qje/qju022.

Elizabeth M. Caucutt, Lance Lochner, and Youngmin Park. Correlation, consumption, confusion, or constraints: Why do poor children perform so poorly? *The Scandinavian Journal of Economics*, 119(1):

102–147, 2017. doi: 10.1111/sjoe.12195. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/sjoe.12195.

Elizabeth M. Caucutt and Lance Lochner. Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy*, 128(3):1065–1147, 2020. doi: 10.1086/704759. URL https://doi.org/10.1086/704759.

Pedro Carneiro, Italo Lopez Garcia, Kjell G. Salvanes, and Emma Tominey. Intergenerational mobility and the timing of parental income. *Journal of Political Economy*, 129(3):757–788, 2021. doi: 10.1086/712443. URL https://doi.org/10.1086/712443.

Daniel Herbst and Nathaniel Hendren. Opportunity unraveled: Private information and the missing markets for financing human capital. Working Paper 29214, National Bureau of Economic Research, September 2021. URL http://www.nber.org/papers/w29214.

Alexander Torgovitsky. Nonparametric inference on state dependence in unemployment. *Econometrica*, 87 (5):1475–1505, 2019. doi: https://doi.org/10.3982/ECTA14138. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA14138.

Martin Nybom and Jan Stuhler. Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation. *Journal of Human Resources*, 51(1):239–268, 01 2016. doi: 10.3368/jhr.51.1.239. URL muse.jhu.edu/article/609071.

Gaëlle Simard-Duplain and Xavier St-Denis. Sample selection in tax data sets of intergenerational links: Evidence from the longitudinal and international study of adults. Research Paper 89-648-x2020002, Statistics Canada, March 2020. URL https://www150.statcan.gc.ca/n1/pub/89-648-x/89-648-x2020002-eng.htm.

Philip Oreopoulos. The Long-Run Consequences of Living in a Poor Neighborhood*. *The Quarterly Journal of Economics*, 118(4):1533–1575, 11 2003. ISSN 0033-5533. doi: 10.1162/003355303322552865. URL https://doi.org/10.1162/003355303322552865.

Marie Connolly, Catherine Haeck, and Jean-William P. Laliberté. *Parental Education and the Rising Transmission of Income between Generations*. University of Chicago Press, February 2021. URL http://www.nber.org/chapters/c14433.

Pierre-André Chiappori, Monica Costa Dias, and Costas Meghir. The marriage market, labor supply, and education choice. *Journal of Political Economy*, 126(S1):S26–S72, 2018. doi: 10.1086/698748. URL https://doi.org/10.1086/698748.

Kim Huynh, Yuri Ostrovsky, Robert Petrunia, and Marcel-Cristian Voia. Industry shutdown rates and permanent layoffs: evidence from firm-worker matched data. *IZA Journal of Labor Economics*, 6, 12 2017. doi: 10.1186/s40172-017-0057-0.

Zhengxi Lin. Employment insurance in canada: Policy changes. Technical Report 75-001-XPE, Statistics Canada, Summer 1998. URL https://www150.statcan.gc.ca/n1/en/pub/75-001-x/1998002/article/3828-eng.pdf.

David M. Gray. Has ei reform unraveled? canada's ei regime in the 2000s. Technical Report 98, C.D. Howe Institute, November 2006. URL https://www.cdhowe.org/sites/default/files/attachments/research_papers/mixed/backgrounder_98.pdf.

Marta Lachowska, Isaac Sorkin, and Stephen A. Woodbury. Firms and unemployment insurance take-up. Conference paper, National Bureau of Economic Research, July 2021. URL https://conference.nber.org/conf_papers/f153580.pdf.

Paul Storer and Marc A. Van Audenrode. Unemployment insurance take-up rates in canada: Facts, determinants, and implications. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 28 (4a):822–835, 1995. ISSN 00084085, 15405982. URL http://www.jstor.org/stable/135933.

Alice O. Nakamura and W. E. Diewert. Insurance for the unemployed: Canadian reforms and their relevance for the united states. Discussion Paper 00-10, University of British Columbia, September 2000. URL https://www.researchgate.net/publication/228549273_Insurance_for_the_unemployed_Canadian_reforms_and_their_relevance_for_the_United_States.

Patricia M. Anderson and Bruce D. Meyer. Unemployment Insurance Takeup Rates and the After-Tax Value of Benefits*. *The Quarterly Journal of Economics*, 112(3):913–937, 08 1997. ISSN 0033-5533. doi: 10.1162/003355397555389. URL https://doi.org/10.1162/003355397555389.

Wayne Vroman. Unemployment insurance recipients and nonrecipients in the cps. *Monthly Labor Review*, 132:44–53, 10 2009. URL https://www.bls.gov/opub/mlr/2009/10/art4full.pdf.

A. Ebenstein and Kevin Stange. Does inconvenience explain low take-up? evidence from unemployment insurance. *Journal of Policy Analysis and Management*, 29:111–136, 2010.

Human Resources Development Canada. Improving social security in canada : a discussion paper. MP90-2/3-1994E, 1994. URL publications.gc.ca/pub?id=9.645681&sl=0.

Michael Baker, Miles Corak, and Andrew Heisz. The labour market dynamics of unemployment rates in canada and the united states. *Canadian Public Policy / Analyse de Politiques*, 24:S72–S89, 1998. ISSN 03170861, 19119917. URL http://www.jstor.org/stable/3551935.

Adam M Lavecchia and Alisa Tazhitdinova. Permanent and transitory responses to capital gains taxes: Evidence from a lifetime exemptionin canada. Working Paper 28514, National Bureau of Economic Research, February 2021. URL http://www.nber.org/papers/w28514.

Peng Zhao, Xiaogang Su, Tingting Ge, and Juanjuan Fan. Propensity score and proximity matching using random forest. *Contemporary Clinical Trials*, 47:85–92, 2016. ISSN 1551-7144. doi: https://doi.org/10.1016/j.cct.2015.12.012. URL https://www.sciencedirect.com/science/article/pii/S1551714415301439.

Susan Athey and Guido Imbens. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1510489113. URL https://www.pnas.org/content/113/27/7353.

Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018. doi: 10.1080/01621459.2017.1319839. URL https://doi.org/10.1080/01621459.2017.1319839.

Susan Athey, Julie Tibshirani, and Stefan Wager. Generalized random forests, 2018.

Rina Friedberg, Julie Tibshirani, Susan Athey, and Stefan Wager. Local linear forests. *Journal of Computational and Graphical Statistics*, 30(2):503–517, 2021. doi: 10.1080/10618600.2020.1831930. URL https://doi.org/10.1080/10618600.2020.1831930.

Susan Athey and Guido W. Imbens. Machine learning methods that economists should know about. *Annual Review of Economics*, 11(1):685–725, 2019. doi: 10.1146/annurev-economics-080217-053433. URL https://doi.org/10.1146/annurev-economics-080217-053433.

Gérard Biau and Erwan Scornet. A random forest guided tour, 2015.

Yi Lin and Yongho Jeon. Random forests and adaptive nearest neighbors. *Journal of the American Statistical Association*, 101(474):578–590, 2006. doi: 10.1198/016214505000001230. URL https://doi.org/10.1198/016214505000001230.

Gérard Biau and Luc Devroye. On the layered nearest neighbour estimate, the bagged nearest neighbour estimate and the random forest method in regression and classification. *Journal of Multivariate Analysis*, 101(10):2499–2518, 2010. ISSN 0047-259X. doi: https://doi.org/10.1016/j.jmva.2010.06.019. URL https://www.sciencedirect.com/science/article/pii/S0047259X10001387.

Matthew A. Olson and Abraham J. Wyner. Making sense of random forest probabilities: a kernel perspective, 2018.

Simon Jager and Jorg Heining. How substitutable are workers? evidence from worker deaths. Working paper, MIT, 2019. URL http://economics.mit.edu/files/16635.

Heather Sarsons. Interpreting signals in the labor market: Evidence from medical referrals. Working paper, University of Chicago, Booth, 2019. URL https://drive.google.com/file/d/12LI5b4Xg7DlNWt-ml2qw-PaMHlihdl0V/view.

Alberto Abadie, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge. When should you adjust standard errors for clustering? Working Paper 24003, National Bureau of Economic Research, November 2017. URL http://www.nber.org/papers/w24003.

V. Chernozhukov, I. Fernàndez-Val, and A. Galichon. Improving point and interval estimators of monotone functions by rearrangement. *Biometrika*, 96(3):559–575, 06 2009. ISSN 0006-3444. doi: 10.1093/biomet/asp030. URL https://doi.org/10.1093/biomet/asp030.

James Uguccioni. Optimal unemployment insurance with dependants. Working paper, University of Toronto, 2020. URL https://www.dropbox.com/sh/41s3cd23nhb8lzw/AABSSu9bnFG3acglUhIt05d8a?dl=0.

Raj Chetty. A general formula for the optimal level of social insurance. *Journal of Public Economics*, 90(10):1879 – 1901, 2006. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2006.01.004. URL http://www.sciencedirect.com/science/article/pii/S0047272706000223.

Michael Baker, Jonathan Gruber, and Kevin Milligan. Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745, 2008. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/10.1086/591908.

Marc Frenette. Postsecondary enrolment by parental income: Recent national and provincial trends. Research Paper 11-626-x2017070, Statistics Canada, April 2017. URL https://www150.statcan.gc.ca/n1/pub/11-626-x/11-626-x2017070-eng.htm.

Jean-William Laliberté. Long-term contextual effects in education: Schools and neighborhoods. *American Economic Journal: Economic Policy*, 13(2):336–77, May 2021. doi: 10.1257/pol.20190257. URL https://www.aeaweb.org/articles?id=10.1257/pol.20190257.

Liberal Party of Canada. Forward. for everyone., 2021. URL https://liberal.ca/wp-content/uploads/sites/292/2021/09/Platform-Forward-For-Everyone.pdf.

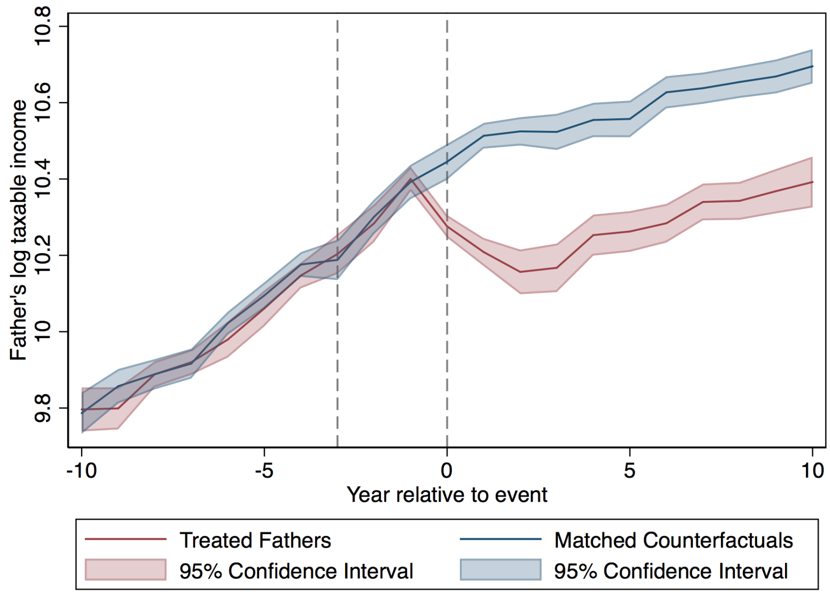
Flavio Cunha and James Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, May 2007. doi: 10.1257/aer.97.2.31. URL https://www.aeaweb.org/articles?id=10.1257/aer.97.2.31.

Gary King, Richard Nielsen, Carter Coberley, James E. Pope, and Aaron Wells. Comparative effectiveness of matching methods for causal inference, 2011.

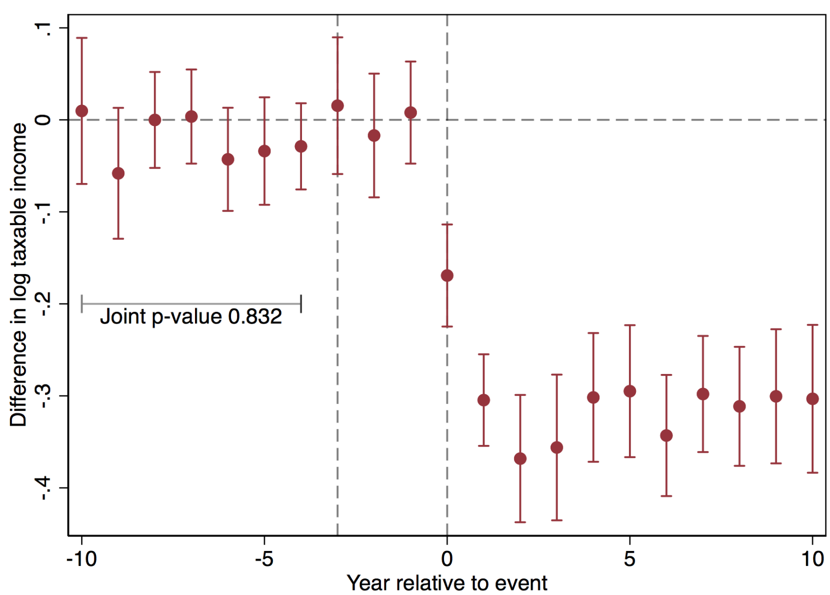
Gary King and Richard Nielsen. Why propensity scores should not be used for matching. *Political Analysis*, 27(4):435–454, 2019. doi: 10.1017/pan.2019.11.

8 Figures

Figure 1: Pretrends for Children Treated at Age 6



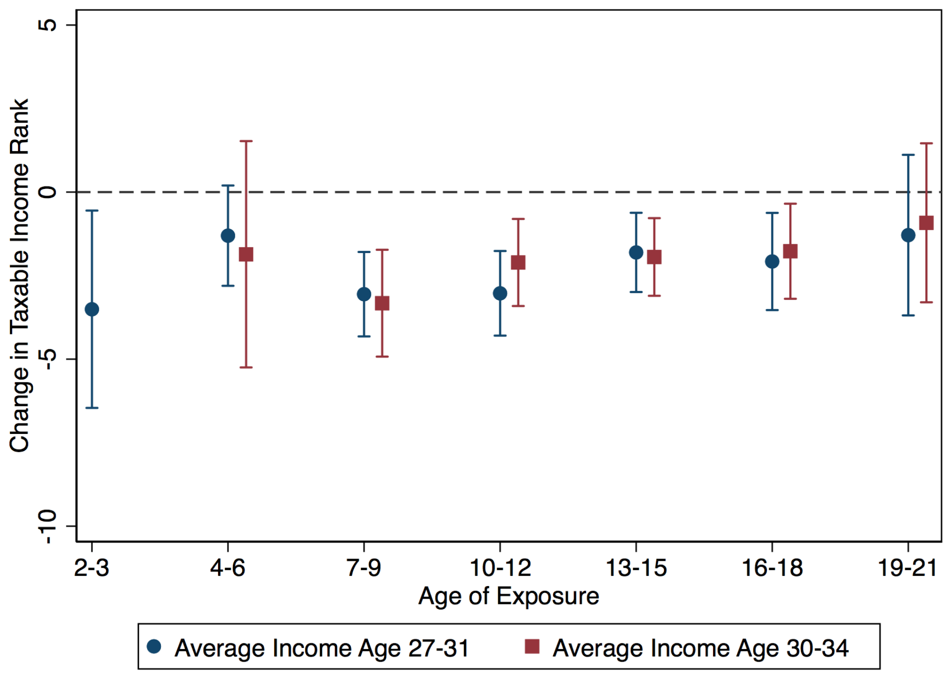
(a) Father's Log Taxable Income in Levels



(b) Difference between Treat and Control Groups

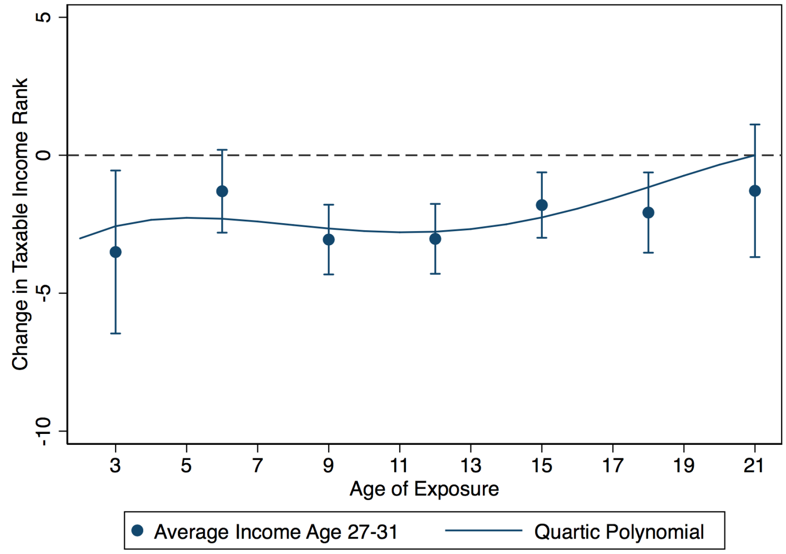
This figure presents the average log taxable income for treated fathers who experience job loss when their child is age 6, alongside their matched counterfactuals. Panel (a) presents average log taxable income in levels, while Panel (b) shows the annual difference between treated and control fathers. The estimates in Panel (b) correspond to δ_t in equation (5), normalized to 0 in year τ. In Panel (b), I also present the p-value testing for balance between by treatment and control fathers, $H_0 : \delta_t = 0 \ \forall t < \tau - 3$. All standard errors are clustered by father/child pair. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Figure 2: Income Attainment Treatment Effects by 3 Year Age Bins

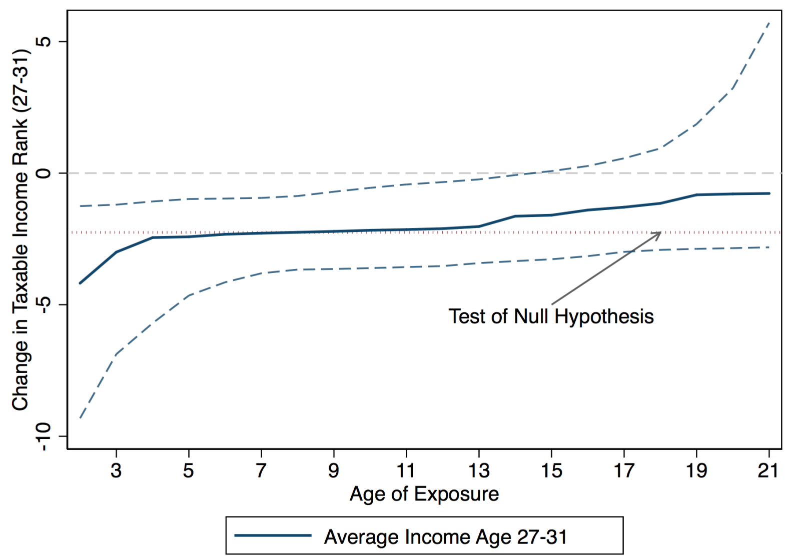


This figure presents the average reduction in a child's taxable income rank from (i) age 27 to 31 and (ii) age 30 to 34 by her age of exposure to parental unemployment. Each point estimate corresponds to an estimate of of Δ_a from the regression presented in equation (4), pooling up to 3 ages of exposure together. In each regression, I control for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \; \forall a$). The error bars on each point estimate are 95% confidence intervals, with standard errors clustered at the match-pair level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Figure 3: Trends in Income Attainment



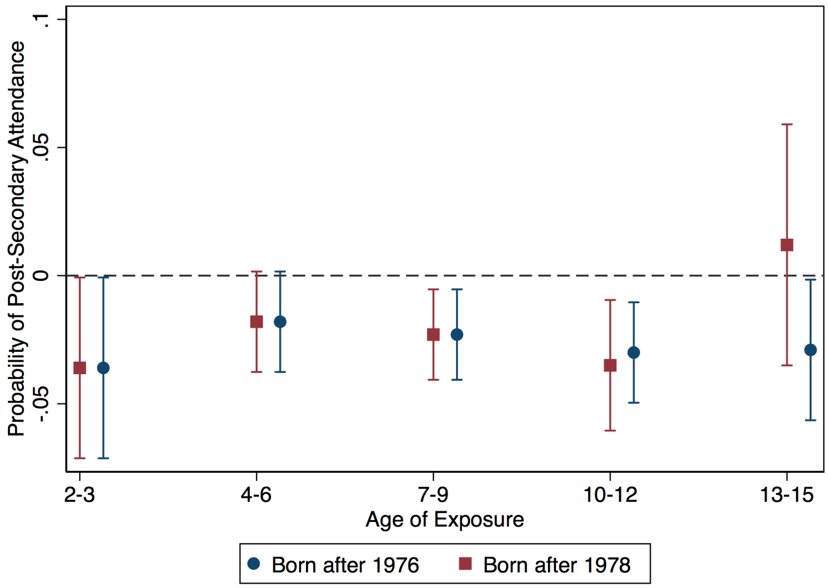
(a) Polynomial Smoothed Reductions in Income Attainment



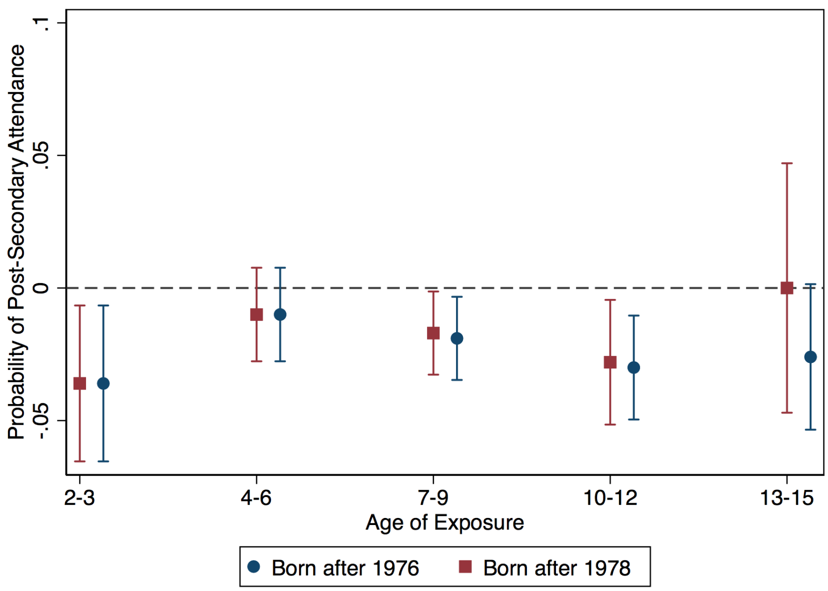
(b) Chernozhukov et al. (2009) test for Monotonicity

Panel (a) presents a smoothed, quartic polynomial of age of exposure to parental unemployment and the implied reduction in a child's income attainment, as measured by their average taxable income rank between the ages of 27 and 31, using the same three year bins as Figure 2. Specifically, I estimate quartic polynomial $Y_i = \sum_{q=0}^{4} \alpha_q a^q + \beta_q D_i(a) a^q + u_i$ and present the fixed values of $\hat{Y} = \sum_{q=0}^{4} \hat{\beta}_q a^q$ for each $a \in \{2, 3, ..., 21\}$. In Panel (b), I test whether there is a monotonic and increasing relationship between the age of exposure to parental unemployment and the reduction in income attainment it causes. I present a nonparametric, local linear regression specification of the trends in my treatment effects by age, constrained to be weakly increasing by rearrangement following Chernozhukov et al. (2009). Alongside the estimated relationship between age of exposure and income attainment losses, I plot a bootstrapped 95% uniform confidence interval obtained from 1,000 replications, with re-sampling stratified at the match-level following Abadie et al. (2017). I visually test the null hypothesis that there is no relationship between age of exposure to parental unemployment and the resulting reduction in the child's income attainment by plotting a horizontal line through the confidence interval. As such, I fail to reject the null hypothesis that the relationship between age of exposure to parental unemployment and a child's income attainment is monotone at the 5% level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Figure 4: Educational Attainment Treatment Effects



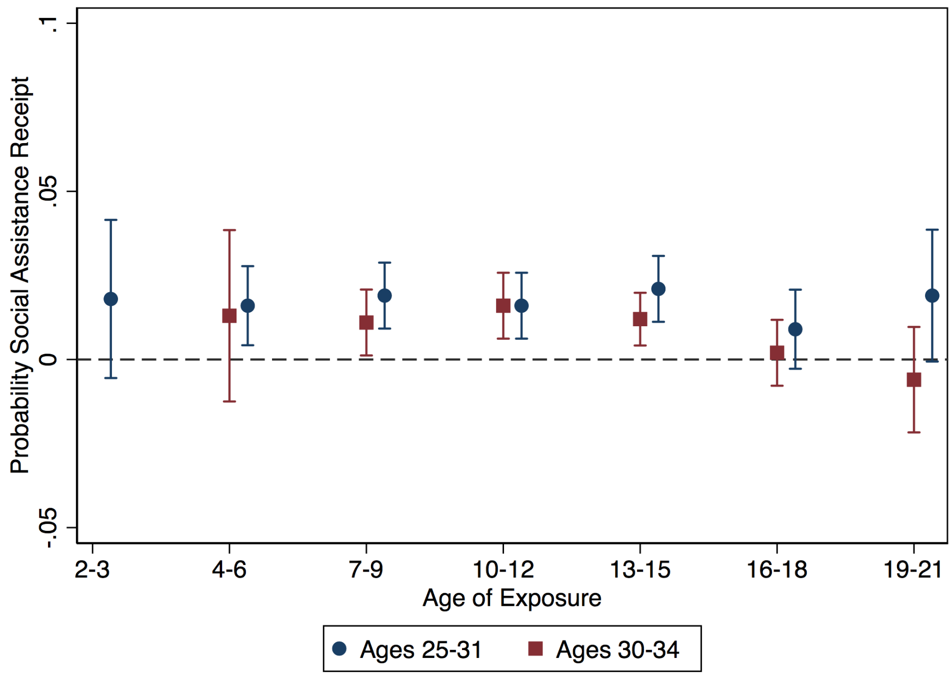
(a) Attending Post-Secondary between Ages 18 and 20



(b) Attending Post-Secondary between Ages 18 and 25

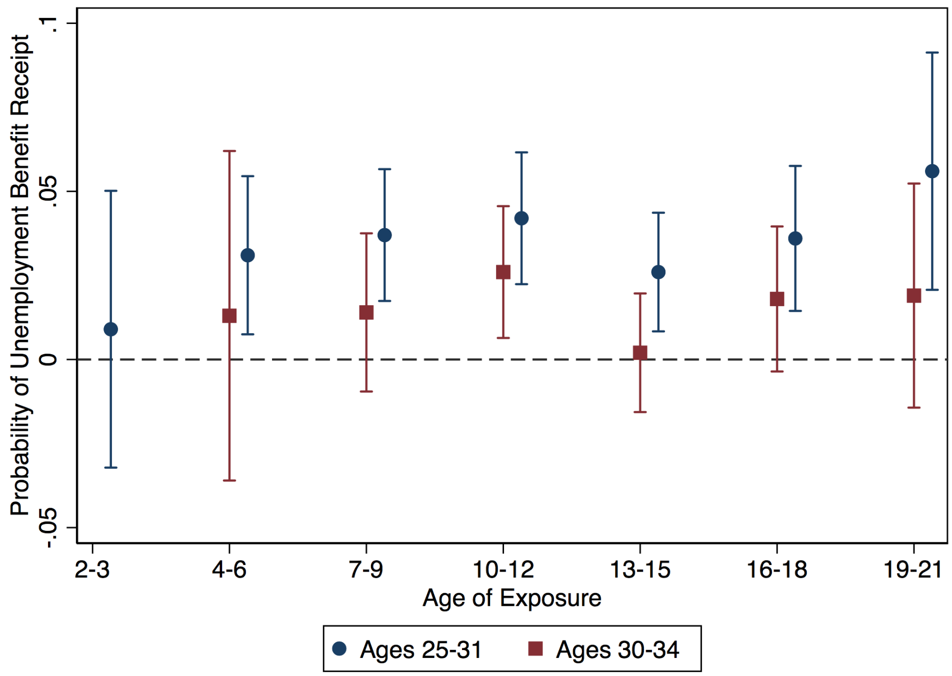
Panels (a) and (b) of this figure respectively present the average reduction in a child's likelihood of claiming post-secondary education tax credits on their tax returns (i) between the ages of 18 and 20 and (ii) between the ages 18 and 25 by her age of exposure to parental unemployment. Each point estimate corresponds to an estimate of of Δ_a from the regression presented in equation (4), pooling up to 3 ages of exposure together. In each regression, I control for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \; \forall a$). The error bars on each point estimate are 95% confidence intervals, with standard errors clustered at the match-pair level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Figure 5: Social Assistance Receipt by Age of Exposure



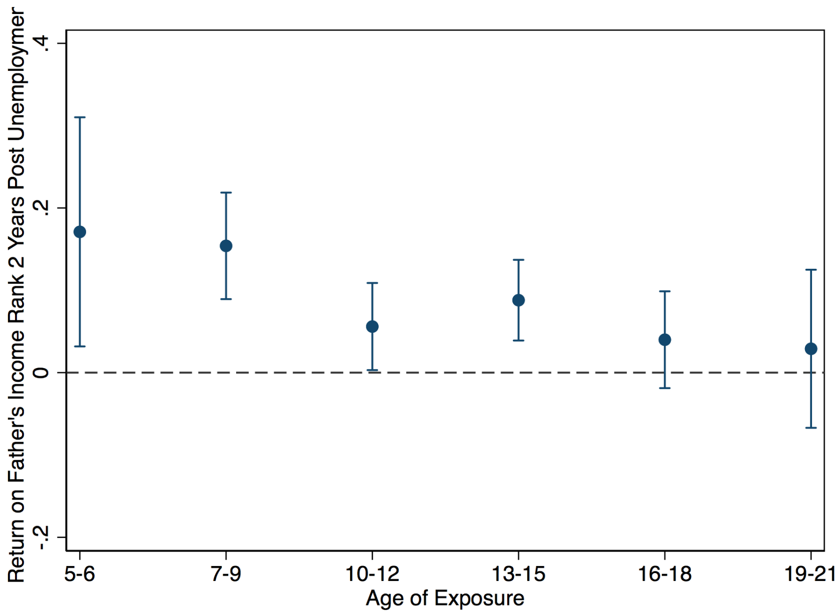
This figure presents the average change in a child's likelihood of receiving social assistance from (i) age 25 to 31 and (ii) age 30 to 34 by her age of exposure to parental unemployment. Each point estimate corresponds to an estimate of of Δ_a from the regression presented in equation (4), pooling up to 3 ages of exposure together. In each regression, I control for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \; \forall a$). The error bars on each point estimate are 95% confidence intervals, with standard errors clustered at the match-pair level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Figure 6: Unemployment Insurance Benefits Receipt by Age of Exposure

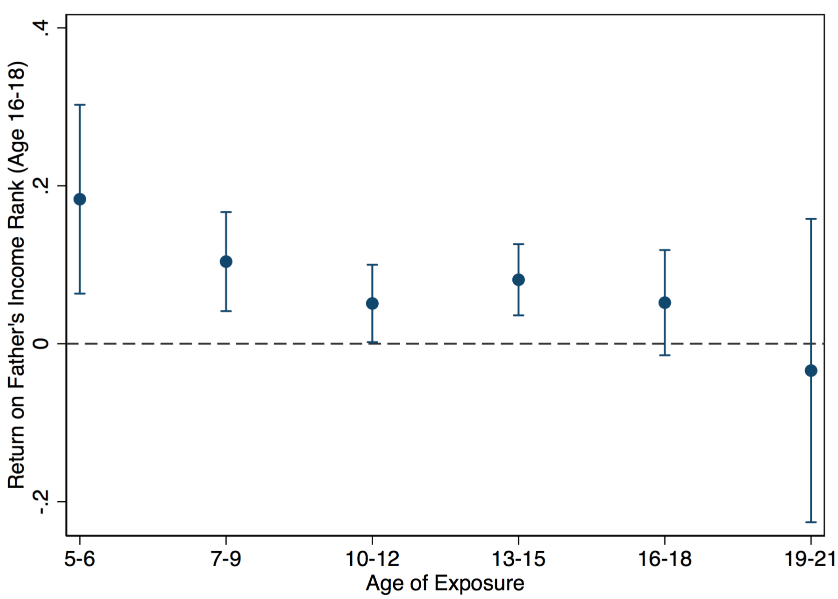


This figure presents the average change in a child's likelihood of receiving unemployment insurance benefits from (i) age 25 to 31 and (ii) age 30 to 34 by her age of exposure to parental unemployment. Each point estimate corresponds to an estimate of of Δ_a from the regression presented in equation (4), pooling up to 3 ages of exposure together. In each regression, I control for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \ \forall a$). The error bars on each point estimate are 95% confidence intervals, with standard errors clustered at the match-pair level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Figure 7: Return on Father's Income - Fixed Window Size



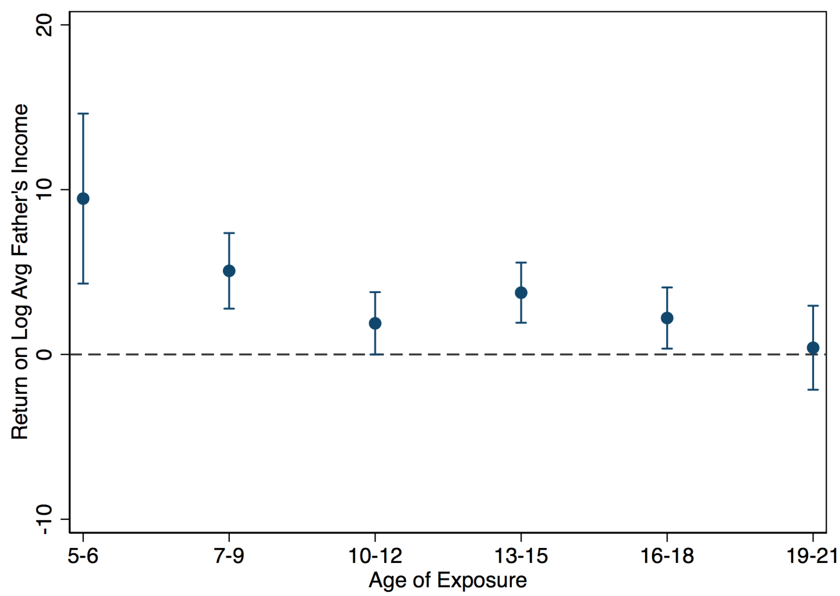
(a) Return on Income Father's Rank in the 2 Years Following Unemployment by 3 Year Age Bins



(b) Return on Income Father's Rank at 16 to 18 by 3 Year Age Bins

This figure presents estimates of $\hat{\psi}_a$ from equation (7) using two different income concepts. In Panel (a), I regress the father's average percentile income rank in the two years after unemployment on the child's income attainment at ages 30 to 34, controlling for match pair fixed effects and the child's sex. In Panel (b), I regress the father's average percentile income rank in when the child is aged 16 to 18 on the child's income attainment at ages 30 to 34, controlling for match pair fixed effects and the child's sex. The error bars on each point estimate are 95% confidence intervals, with standard errors clustered at the match-pair level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Figure 8: Return on Father's Scarring Income by 3 Year Age Bins



This figure presents estimates of $\hat{\psi}_a$ from equation (7) using the father's log total taxable income between the year of initial job loss and the year the child reaches age 21. These correspond to the estimates of $\hat{\psi}_a$ in Table 6, and are conditional on the a dummy variable for the father experiencing divorce after unemployment, the average educational attainment of the father's Census Sub-Division of residence after unemployment, match pair fixed effects, and the child's sex. The error bars on each point estimate are 95% confidence intervals, with standard errors clustered at the match-pair level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

9 Tables

Table 1: Match-Year Summary Statistics

	Employment Status		Received UI Benefits	
	Unemployed	Employed	Yes	No
	(1)	(2)	(3)	(4)
Household				
Father married	0.915	0.946	0.841	0.893
	(0.000)	(0.000)	(0.000)	(0.000)
Father's spouse in Labour Force	0.644	0.644	0.753	0.744
	(0.000)	(0.000)	(0.000)	(0.000)
Number of children	1.89	1.93	2.66	2.64
	(0.00)	(0.00)	(0.00)	(0.00)
Province				
Atlantic	0.147	0.094	0.092	0.077
	(0.000)	(0.000)	(0.000)	(0.000)
Québec	0.304	0.262	0.286	0.240
	(0.000)	(0.000)	(0.000)	(0.000)
Ontario	0.237	0.365	0.359	0.377
	(0.000)	(0.000)	(0.000)	(0.000)
Prairies	0.167	0.173	0.164	0.199
	(0.000)	(0.000)	(0.000)	(0.000)
British Columbia	0.145	0.107	0.094	0.101
	(0.000)	(0.000)	(0.000)	(0.000)
N	95,859	1,449,801	1,310,200	13,988,900

Notes: This table presents means and standard errors (in brackets) for a number of observables from match years 1984-1990 for treatment years 1987-1993. Columns (1) and (2) report statistics in the match year for all fathers in the Labour Force Survey. Columns (3) and (4) report statistics in the match year for all tax filing fathers in the IID who respectively did and did not receive UI benefits in the treatment year. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Table 2: Match-Year Summary Statistics

	Received UI Benefits		Possible to Match		Matched Sample	
	Yes	No	Treated	Control	Treated	Control
	(1)	(2)	(3)	(4)	(5)	(6)
Father						
Net taxable income	26000	34600	27400	34800	32700	32500
	(13)	(9)	(20)	(10)	(107)	(107)
Spouse on tax return	0.841	0.893	0.855	0.895	0.972	0.98
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.00)
Father's Spouse						
Net taxable income	11900	14600	11900	14400	15600	15800
	(163)	(10)	(381)	(11)	(90)	(83)
Frac positive labour income	0.753	0.744	0.753	0.743	0.631	0.629
	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.003)
Household						
Net taxable income	32900	44400	34600	44400	44100	44000
	(100)	(12)	(240)	(13)	(145)	(140)
Number of children	2.66	2.64	2.64	2.64	2.46	2.44
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Child						
Sex (female=1)	0.456	0.469	0.459	0.469	0.466	0.473
	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.003)
Father's age at birth	29.4	30.6	29.7	30.7	30.1	30.0
	(0.01)	(0.00)	(0.01)	(0.00)	(0.03)	(0.03)
Mother's age at birth	26.5	27.7	26.8	27.8	27.7	27.7
	(0.01)	(0.00)	(0.01)	(0.00)	(0.03)	(0.03)
Province						
Atlantic	0.092	0.077	0.078	0.078	0.039	0.037
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Québec	0.286	0.240	0.269	0.240	0.298	0.302
	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.003)
Ontario	0.359	0.377	0.391	0.377	0.441	0.446
	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.003)
Prairies	0.164	0.199	0.163	0.200	0.152	0.148
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.002)
British Columbia	0.094	0.101	0.093	0.101	0.069	0.066
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
N	1,310,200	13,988,900	550,200	13,235,100	31,200	31,200

Notes: This table presents means and standard errors (in brackets) for a number of observables from match years 1984-1990 for treatment years 1987-1993. All income variables are in 1990 Canadian dollars. Columns (1) and (2) report statistics in the match year for all tax filing fathers in the IID who respectively did and did not receive UI benefits in the treatment year. Columns (3) and (4) report statistics in the match year for all treated and control observations who could possibly be matched respectively. Columns (5) and (6) report report statistics in the match year for all treated and control observations in my preferred matched sample respectively (that is, all observations matched with a proximity greater than 0.7). All estimates are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Table 3: Effects of Parental Unemployment on Child's Income Attainment

	Income Rank Age 27 to 31	**Income Rank Age 30 to 34**
Age of Exposure	(1)	(2)
2-3	-3.746^{**}	
	(1.513)	
4-6	-1.326^{*}	-1.804
	(0.770)	(1.735)
7-9	-3.110^{***}	-3.271^{***}
	(0.647)	(0.820)
10-12	-3.119^{***}	-2.191^{***}
	(0.649)	(0.668)
13-15	-1.920^{***}	-2.034^{***}
	(0.607)	(0.594)
16-18	-2.117^{***}	-1.826^{**}
	(0.744)	(0.728)
19-21	-1.318	-0.896
	(1.234)	(1.222)
N	37,500	28,700
Mean Control Outcome	53.9	53.4

Notes: This table presents point estimates of Δ_a from the regression presented in equation (4), controlling for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \ \forall a$). Standard errors presented in parentheses are clustered at the match-pair level. These estimates are also presented in Figure 2. All estimates are for 3 year age bins, with two exceptions: (i) in Column 1, the first row only covers children aged 2 to 3 at exposure to parental job loss; and (ii) in Column 2, the first non-empty row only covers children aged 5 to 6 at exposure to parental job loss. *, **, and *** respectively indicate statistical significance at the 10%, 5%, and 1% levels. All estimates are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Table 4: Childhood Pretrends

	Father's Income			**Father's Spouse's Income**			**Household**		**Filed Taxes**		
Age	Taxable	Labour	Net of Tax	Taxable	Labour	Net of Tax	Married	Moved CSD	Father	Father's Spouse	Father's Spouse T4
2	0.457	0.256	0.409	0.960	0.816	0.949	0.999	0.457	0.878	0.608	0.500
3	0.670	0.428	0.651	0.947	0.885	0.947	0.438	0.186	0.324	0.175	0.404
4	0.271	0.722	0.223	0.026	0.076	0.023	0.877	0.033	0.137	0.362	0.530
5	0.470	0.941	0.498	0.542	0.772	0.544	0.549	0.596	0.553	0.644	0.716
6	0.832	0.650	0.902	0.447	0.610	0.451	0.920	0.521	0.701	0.698	0.489
7	0.369	0.783	0.228	0.543	0.565	0.567	0.511	0.882	0.776	0.856	0.391
8	0.357	0.943	0.300	0.539	0.340	0.544	0.815	0.817	0.738	0.534	0.616
9	0.182	0.102	0.138	0.812	0.874	0.818	0.326	0.240	0.601	0.299	0.389
10	0.119	0.936	0.084	0.409	0.294	0.402	0.557	0.639	0.592	0.665	0.373
11	0.189	0.866	0.105	0.937	0.796	0.943	0.957	0.135	0.940	0.695	0.553
12	0.114	0.217	0.118	0.540	0.603	0.601	0.965	0.610	0.924	0.640	0.262
13	0.624	0.269	0.589	0.319	0.890	0.292	0.987	0.300	0.092	0.993	0.692
14	0.727	0.947	0.653	0.605	0.612	0.587	0.254	0.794	0.942	0.994	0.735
15	0.166	0.891	0.225	0.399	0.576	0.412	0.514	0.701	0.797	0.905	0.039
16	0.271	0.798	0.119	0.874	0.955	0.866	0.591	0.659	0.341	0.850	0.595
17	0.246	0.982	0.215	0.402	0.557	0.452	0.335	0.116	0.303	0.951	0.609
18	0.043	0.931	0.038	0.287	0.728	0.294	0.989	0.763	0.630	0.333	0.826
19	0.471	0.112	0.383	0.761	0.713	0.771	0.409	0.122	0.103	0.677	0.851
20	0.872	0.854	0.729	0.029	0.073	0.028	0.174	0.044	0.943	0.479	0.301
21	0.030	0.081	0.024	0.323	0.333	0.317	0.684	0.734	0.544	0.744	0.176

Notes: These are the resulting p-values from testing the $\delta_t \; \forall t < \tau - 3$ in equation (5) for a given outcome. Each column presents a different outcome, and each row presents age of exposure to parental unemployment for the matched treatment and control sample. Standard errors are clustered at the level of child i. Labour income is measured on the T4 form, and excludes self-employment income as a result. The "Married" outcome includes married or co-habitating partners. Measures of whether father or father's spouse filed taxes are dummy variables for whether taxes records are available in a given year. Similarly, "Father's Spouse T4" is a dummy variable for whether the father's spouse received any labour income in a given tax year, measured on the T4 form.

Table 5: Additional SES Outcomes

	Marital Status		Children	
	Age 25	Age 30	Any by 30	Age at First Birth
	(1)	(2)	(3)	(4)
Age 2-3	-0.036	-0.009	-0.009	-0.407
	(0.015)	(0.020)	(0.014)	(0.601)
Age 4-6	-0.005	-0.008	0.021	-0.469
	(0.008)	(0.011)	(0.008)	(0.328)
Age 7-9	0.000	-0.009	0.010	-0.049
	(0.007)	(0.009)	(0.006)	(0.249)
Age 10-12	-0.003	0.000	0.003	-0.388
	(0.007)	(0.009)	(0.006)	(0.245)
Age 13-15	-0.008	-0.015	0.009	-0.223
	(0.006)	(0.008)	(0.006)	(0.209)
Age 16-18	0.002	0.002	0.000	0.026
	(0.007)	(0.010)	(0.006)	(0.240)
Age 19-21	-0.019	-0.025	0.012	-0.432
	(0.012)	(0.016)	0.010	(0.237)
N	28,900	28,900	28,900	3050
Control Group Mean	0.126	0.282	0.100	29.6

Notes: This table presents point estimates from (4) for $\hat{\Delta}_a$ for various outcomes. Columns 1, 2, and 3 use dummy variables respectively indicating whether a child was married at age 25, married at age 30, and had list any dependants on her tax return by age 30. Column 4 conditions on having had a dependant, and considers the average age of treated and control children when they first list this dependant on their tax return. All standard errors are clustered at the match pair level, following Abadie et al. (2017). All estimates are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Table 6: Mediation Analysis Decomposition

	Income			**Neighbourhood**			**Divorce**			% Δ_a
	Change	Slope	Pct	Change	Slope	Pct	Change	Slope	Pct	Explained
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age 5-6	-0.233	9.37	122%	-0.0044	13.9	3%	0.013	-0.037	0%	126%
Age 7-9	-0.274	4.99	41%	-0.0092	9.62	3%	0.027	-3.787	3%	47%
Age 10-12	-0.249	1.94	22%	-0.0050	19.6	5%	0.025	0.818	-1%	26%
Age 13-15	-0.256	3.79	49%	-0.0037	20.8	4%	0.013	-2.633	2%	55%
Age 16-18	-0.297	2.21	36%	-0.0132	1.13	1%	0.002	0.885	0%	37%
Age 19-21	-0.233	0.52	12%	-0.0049	28 .3	14%	0.030	4.113	-12%	14%

Notes: This table presents a mediation analysis considering how (i) log father's reduced income, (ii) changes to neighbourhood quality, and (iii) changes in the incidence of divorce following an unemployment spell affect children. For each input, I present estimates of the change in an input, the slope (return) relating that input to the child's income attainment at age 30 to 34, and the percentage of the overall treatment effect that that input explains. In particular, in Columns (1), (4), and (7), I present estimates of the average reduction in a given input for children first exposed to parental unemployment at age a, $\hat{\phi}_a$ from equation (6). In Columns (2), (5), and (8), I present estimates of the return on a given input for children first exposed to parental unemployment at age a, $\hat{\psi}_a$ from equation (7). Columns (3), (6), and (9) are the product of the two preceding columns divided by the estimated reduction in income attainment at ages 30 to 34 for children first exposed at age a. Column (10) sums Columns (3), (6), and (9) and presents how much of my estimated effect from Table 3 is explained by these three inputs. All estimates are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

A1 Model Appendix

In this Section, I present a simple model which motivates the potential outcomes and the treatment effects presented in Section 2. Suppose parents live for 2 periods, and are endowed with a child which they raise throughout their adulthood. In each period, parents realize their income/employment status and allocate their resources among savings, human capital investments in their children, and consumption. This model explores how child i's income attainment is affected by the decisions her parent $j(i)$ makes throughout her childhood. Specifically, let child i's adulthood human capital stock be,

$$Y_i = \left(\gamma_1 i_{1i}^{\theta} + \gamma_2 i_{2i}^{\theta}\right)^{\frac{1}{\theta}}$$

where i_t is an investment made in the child at age a, and γ_a and θ governs the importance and substitutability of investments made in periods 1 to 2. For simplicity, assume that the return on the labour market for each unit of human capital is \$1 when the child reaches adulthood, so human capital attainment is equal to adulthood earnings attainment.

Parents solve their optimal human capital investment problem sequentially. Suppose that parent $j(i)$'s earnings $Z_{j(i)a}$ in period a are known given her employment status, D_a,

$$Z_{j(i)a} = (1 - D_a)Z_{j(i)a}(0) + D_a Z_{j(i)a}(1)$$

with employment status D_a being revealed at the beginning of each period. In every period a, assume unemployed earnings are strictly less than employed earnings, $Z_{j(i)a}(1) < Z_{j(i)a}(0)$, but that they are potentially non-zero as a result of unemployment insurance. Assume that a parent transitions between unemployment according to her potential outcomes $\{U_a(0), U_a(1)\}_{a\in\{1,2\}}$, as introduced in Section 2. In this simple example, wage scarring from unemployment only enters the model through transition probabilities, e.g. $Pr(D_{a+1} = 1|D_a = 1) > Pr(D_{a+1} = 1|D_a = 0)$[39].

Parents receive a bequest B^P of wealth in the first period, then choose consumption c_a, savings s_a and human capital investments i_a in every period to solve,

Period 1

$$\begin{aligned}
V_1(B^P; D_1) \quad &= \quad \max_{c_1, i_1, s_1} \{u(c_1) + \beta \mathbb{E}_{D_2}\left[V_2(s_1, i_1; D_2)|D_1\right]\} \\
& \quad\quad c_1 + p_1 i_1 + \frac{1}{R} s_1 \leq Z_1(D_1) + B^P \\
& \quad\quad s_1 \geq -m_1
\end{aligned}$$

[39] I relax this assumption in the empirical application, allowing for any wage scarring which is the result of an unemployment shock.

Period 2

$$
\begin{aligned}
V_2(s_1, i_1; D_2) \quad = \quad & \max_{c_2, i_2, s_2} \left\{ u(c_2) + \delta U^C(Y_i + B_C) \right\} \\
& c_2 + p_2 i_2 + \frac{1}{R} B^C \leq Z_2(d_2) + s_1 \\
& B^C \geq 0
\end{aligned}
$$

where B^C is the parent's monetary bequest given to her children, m_1 is a borrowing constraint imposed on all parents, and β and δ are time and benevolence preference parameters which respectively discount the future and the utility of the child.

In this model, unemployment can distort the human capital investments made by parents through the borrowing constraints, and affect children in the long-run. In each period, the parent's decision ultimately comes down to weighing the marginal value of investing in their child against the return on a dollar of savings. The intuition being, if a dollar of savings has a higher return than investing another dollar in human capital, then parents are better off saving the dollar and transferring it to their child in adulthood. In particular, in period 2 this amounts to,

$$
\frac{\lambda_1}{\frac{1}{R}\lambda_1 - \mu_1} p_1 = \frac{\partial Y_i}{\partial i_2}
$$

where λ_2 and μ_2 are respectively the shadow prices on the budget and borrowing constraints in period 2. Crucially, notice that when $\mu_2 = 0$ an additional dollar of earnings will not affect the human capital investments made by parents and be allocated between consumption and the bequest. As such, all parents who are not borrowing constrained in period 2 will make the same investment decision, driving the return on human capital investment down to the risk free rate. In period 1, an additional dollar of earnings today can increase human capital investments even when parents are not borrowing constrained in the present period. In particular, if human capital investments are not perfectly substitutable between periods and parents anticipate they could be borrowing constrained with positive probability in the future, then $\frac{\partial i_a}{\partial Z_j(i)a} \neq 0$ [40].

This simple model provides several important predictions when considering how parental unemployment can affect children later on in life. First off, if parents are not borrowing constrained during their lifetimes or they were allowed to borrow against the future earnings of their children to make human capital investments, then there would be no differences in the human capital attainment across children whose parents realized different earnings during their lifetime[41]. Only imperfect credit markets can create a wedge between the return on savings and on a dollar of human capital investment, leaving equally skilled children with different earnings attainment in adulthood.

[40] This follows from the proof of this property in Caucutt and Lochner (2020).

[41] Allowing for parents and children to have correlated skills would negate this result, as shown in much of the intergenerational literature.

Second, if borrowing constraints may bind in a future period, parents will underinvest in their children in the present. In this case, parents who could be borrowing constrained anticipate a lower expected return on investments made in the present due to complementarities, and reach an expected return equal to the risk free rate at a lower level of i_a. Naturally, this investment gap between unconstrained and possibly constrained parents is increasing in the likelihood of a parent's future unemployment. As such, two children whose parents face different risks of future unemployment may receive different investments throughout childhood and attain different levels of income in adulthood, even in the event neither experienced parental unemployment nor lived with a borrowing constrained parent.

Finally, underinvestment in children in any past period can lead to underinvestment in every period whenever $\theta < 1$. Though past unemployment does not predict future unemployment given that a parent is employed in the present, complementarities in human capital investment still cause past unemployment to affect the child's income attainment. These complementarities result in reduced returns on human capital investments made at any age, relative to the return on the same level of investment made in a child whose parents never faced borrowing constraints.

Deriving the investment elasticity To derive $\frac{\partial i_2}{\partial i_1}\frac{i_1}{i_2}$, I assume that parents are only borrowing constrained in the period of the shock (period 1 in this case) and are not constrained in period 2. When this is the case, the FOC for investing in period 2 is given by

$$\begin{aligned} Rp_2 &= \frac{\gamma_2 i_2^{\theta-1}}{\gamma_1 i_1^{\theta} + \gamma_2 i_2^{\theta}} \\ Rp_2\gamma_1 i_1^{\theta} &= \gamma_2 i_2^{\theta-1} - Rp_2\gamma_2 i_2^{\theta} \end{aligned}$$

Differentiating both sides with respect to i_1,

$$\begin{aligned} Rp_2\theta\gamma_1 i_1^{\theta-1} &= \left((\theta-1)\gamma_2 i_2^{\theta-2} - Rp_2\theta\gamma_2 i_2^{\theta-1}\right)\frac{\partial i_2}{\partial i_1} \\ Rp_2\theta\frac{\partial Y}{\partial i_1} &= \left(\frac{(\theta-1)\gamma_2 i_2^{\theta-2}}{\gamma_1 i_1^{\theta} + \gamma_2 i_2^{\theta}} - Rp_2\theta\frac{\partial Y}{\partial i_2}\right)\frac{\partial i_2}{\partial i_1} \\ Rp_2\theta\frac{\partial Y}{\partial i_1} &= \left((\theta-1)\ i_2^{-1}Rp_2 - Rp_2\theta\frac{\partial Y}{\partial i_2}\right)\frac{\partial i_2}{\partial i_1} \\ \theta\frac{\partial Y}{\partial i_1}i_1 &= \left((\theta-1)\ -\theta\frac{\partial Y}{\partial i_2}i_2\right)\frac{\partial i_2}{\partial i_1}\frac{i_1}{i_2} \\ \frac{\partial i_2}{\partial i_1}\frac{i_1}{i_2} &= \frac{\eta_1}{1-\frac{1}{\theta}-\eta_2} \end{aligned}$$

A1.1 Decomposing the Effects of Parental Unemployment on Children

Using this model, we can express the income attainment of a given child i who experiences the parental employment history (D_1, D_2) in terms of the timing of and the duration of exposure to parental unemployment. To understand the various channels through which gaps in income attainment can arise, it is informative to compare child i with a counterfactual child i'. Child i' has parents $j(i')$ such that $Z_{j(i')a}(D_a) = Z_{j(i)a}(D_a)\ \ \forall a$ and risk of unemployment in every period $j(i)$ is employed is the same as $j(i')$. Nonetheless, i' ultimate experiences parental employment history $(0, 0)$.

Consider the income attainment $Y_a = Y(0, 1)$ of child a whose parent realized employment history $(0, 1)$. By taking a first order Taylor expansion, we can express the gap log income attainment between child a and a' in terms of elasticities which arise from the model,

$$\ln(Y(0,1)) - \ln(Y(0,0)) = \eta_2 \frac{i_{a2} - i_{a'2}}{i_{a'2}}$$

where η_2 is the output elasticity of investments made in period 3[42]. The interpretation of this particular decomposition straight forward — child a will experience a reduction in earnings attainment if her parents are borrowing constrained, $i_{a2} < i_{a'2}$.

From the point of view of period 3, all prior investments were fixed, so complementarities human capital investments do not affect the scale of the loss a experiences. An alternative parental employment history where complementarities may affect the child's overall losses is $(1, 0)$, experienced by child b, though the unemployment spell again only lasts a single period. As with children a and a', consider comparing child b with a counterfactual child b' who experiences parental employment history $(0, 0)$. Using the groups described in Section 2, we should think of parents $j(b)$ and $j(b')$ as having potential outcomes $U_2(0) = U_2(1) = 0$. Unlike in the case of child a, child b's parents may be borrowing constrained in the second period, thereby affecting i_3 as well. In particular, we can decompose total loss in earnings attainment into,

$$\begin{aligned}
\ln(Y(1,0)) - \ln(Y(0,0)) &= (i_{b1} - i_{b'1})\frac{\partial Y_{b'}}{\partial i_1} + (i_{b1} - i_{b'1})\frac{\partial Y_{b'}}{\partial i_2}\frac{\partial i_2}{\partial i_1} \\
&= \eta_1 \frac{i_{b1} - i_{b'1}}{i_{b'1}} + \eta_2 \frac{i_{b1} - i_{b'1}}{i_{b'1}} \frac{\partial i_2}{\partial i_1}\frac{i_{b'1}}{i_{b'2}} \\
&= \eta_1 \frac{i_{b1} - i_{b'1}}{i_{b'1}} + \eta_2 \frac{i_{b1} - i_{b'1}}{i_{b'1}} \frac{\eta_1}{1 - \eta_2 - \frac{1}{\theta}}
\end{aligned}$$

where the first term captures the importance of investments made in period 2 and the second term captures the degree to which investments in the third period are affected by those in the second. For intuition, it is helpful to note that as $\theta \to 1$ the CES function approaches perfect substitutes. In this case, the second term approaches $-\eta_1\%\Delta i_1$ and as a result parents are able to completely compensate for any forgone investments

[42] In particular, $\eta_t = \frac{\gamma_t i^{\theta}_{a't}}{\gamma_1 i^{\theta}_{a'1} + \gamma_2 i^{\theta}_{a'2}}$

from the first period, resulting in $\ln(Y(1,0)) - \ln(Y(0,0)) \to 0$.

Focussing on the spillover of lost investments in period 1 into period 2, there are larger distortions when: human capital is more elastic to i_1; human capital is more elastic to i_2; and the investments in human capital are more complementary[43]. The potential for spillovers in this setting highlights how the timing of an unemployment spell may affect the total loss in human capital for the child. Even if $\eta_1 = \eta_2$ and $\%\Delta i_1 = \%\Delta i_2$, spillovers of past investments mean the child who experiences (1,0) will attain lower human capital than the child who experiences (0,1) despite both children experiencing a single period of parental unemployment. Contrarily, when $\eta_1 < \eta_2$ due to critical periods of human capital investment, it is entirely possible that the child who experiences (1,0) will attain higher human capital than the child who experiences (0,1).

In any event, comparing the effects of parental employment histories (1,0) and (0,1) is a question of timing. Both parental unemployment spells last a single period, and understanding which has the gravest effect on children is the sum of critical periods for investment and dynamic spillovers. The extensive literature studying the scarring effects of unemployment has documented that one of the main sources of lost earnings is a higher incidence of future unemployment following an initial layoff. As such, it is also important to consider how the length of exposure to parental unemployment affects children[44].

As in the cases above, consider child c who experiences parental employment history (1,1) and a counterfactual child c' whose parent is employed in every period. Using the notation from Section 2, parents $j(c)$ and $j(c')$ exhibit positive state dependence and have potential outcomes $U_2(0) = 0$, $U_2(1) = 1$. Further, suppose all of the children who have been considered so far have parents with the exact same earnings potential. Consequently, notice that for the counterfactual children we have $Y_{a'} = Y_{b'} = Y_{c'} = Y(0,0)$. We can express the total loss in earnings for child c as,

$$\begin{aligned}
\ln(Y(1,1)) - \ln(Y(0,0)) &= \ln(Y_c) - \ln(Y_b) + \ln(Y_b) - \ln(Y_{c'}) \\
&= \frac{i_{c2} - i_{b2}}{i_{b2}} \frac{\gamma_2 i_{b2}^{\theta}}{\gamma_1 i_{b1}^{\theta} + \gamma_2 i_{b2}^{\theta}} + \eta_2 \frac{i_{c1} - i_{c'1}}{i_{c'1}} + \eta_3 \frac{i_{c1} - i_{c'1}}{i_{c'1}} \frac{\eta_1}{1 - \eta_2 - \frac{1}{\theta}} \\
&= \frac{i_{c2} - i_{b2}}{i_{b2}} \frac{\gamma_2 i_{b2}^{\theta}}{\gamma_1 i_{b1}^{\theta} + \gamma_2 i_{b2}^{\theta}} + \ln(Y_b) - \ln(Y_{b'})
\end{aligned}$$

Unsurprisingly, the additional period of parental unemployment means child c experiences a weakly larger

[43]In particular, it is helpful to note that $\frac{1}{\theta} = \frac{\sigma}{1-\sigma} =$ where σ is the elasticity of substitution between any two inputs in the CES production function. When $\sigma \in (0,1)$ inputs are complements and lower values of σ are associated with i_2 being more elastic to i_1. The same applies when $\sigma \in (1, \infty)$ and inputs are substitutes — the lower σ and the less substitutable two inputs are, the more elastic i_2 is to i_1. Of course, there the expression of elasticity of i_2 with respect to i_1 implies an asymptote in the Cobb-Douglas case where $\theta = 0$, however it can be shown that $\frac{\partial i_2}{\partial i_1} \frac{i_1}{i_2} = \frac{\gamma_1}{1-\gamma_2} \equiv \frac{\eta_1}{1-\eta_2}$ in that case.

[44]Note that the transitions into and out of unemployment are taken as exogenous in this model, but thiscould be weakened to assuming search effort is independent of the child's age without changing the results in this section of the paper. However, assuming search does not vary with the child's age is inconsistent with forward-looking parents. Indeed, if parents are aware of the dynamic spillovers which result from unemployment in a single period, then their search behaviour should vary systematically with the age of their child. In other work (Uguccioni, 2020), I consider how these dynamic spillovers affect search decisions and optimal unemployment insurance, as they not consistent with Assumption 5(c) in Chetty (2006).

human capital loss than child b. In the event that borrowing constraints bind for the parent of child c in period 2, then $i_{c2} < i_{b2}$ and first term governs how much larger the human capital losses experienced by child c are compared to child b. While first term on the final line looks similar to $\ln(Y_a) - \ln(Y_{a'})$, there are important differences between the two. First, the relevant counterfactual investment lost in period 2 is i_{b2} rather than $i_{c'2} = i_{a'2}$, as underinvestment in period 1 means any investments in period 2 will pay a lower return. Second, it is ambiguous whether $\frac{\gamma_2 i_{b2}^{\theta}}{\gamma_1 i_{b1}^{\theta} + \gamma_2 i_{b2}^{\theta}}$ is higher or lower than η_2 because diminishing marginal returns and dynamic complementarities in investment are working in opposite directions. In particular, diminishing marginal returns on investments means the average return on a dollar should be lower for the child whose (counterfactual) received more, driving towards $\frac{\gamma_2 i_{b2}^{\theta}}{\gamma_1 i_{b1}^{\theta} + \gamma_2 i_{b2}^{\theta}} > \eta_2$. If $\theta < 0$, there are dynamic complementarities in human capital investments and forgone past investments such as $i_{b1} < i_{a'1}$ will decrease the return on investments in period , pushing in the direction of $\frac{\gamma_2 i_{b2}^{\theta}}{\gamma_1 i_{b1}^{\theta} + \gamma_2 i_{b2}^{\theta}} < \eta_2$.

The human capital losses experienced by children a, b, and c are illustrative in explaining how parental unemployment can affect children. Comparing the losses experienced by children a and b, it is clear that the timing of unemployment can explain which child is worse off. For example, suppose that both children experience the same drop in human capital investments due to parental unemployment, $\%\Delta i_1 = \%\Delta i_1 = 1$. In this case, the timing of unemployment is worse for child a if,

$$
\begin{aligned}
\ln(Y(1,0)) - \ln(Y(0,0)) \quad &> \quad \ln(Y(0,1)) - \ln(Y(0,0)) \\
\eta_2 \quad &> \quad \eta_1 + \eta_2 \frac{\eta_1}{1 - \eta_2 - \frac{1}{\theta}}
\end{aligned}
$$

In other words, the timing of parental unemployment is worse in period 2 than 1 if forgone investments made in period 2 have a larger effect on human capital attainment than the sum of (i) forgone investments in period 1 and (ii) spillovers of forgone investments into period 2. In the context of the human capital literature (e.g. Cunha and Heckman, 2007), the η_t terms capture the relative importance of critical periods for investments, while the second term in the losses experienced by b captures dynamic complementarities in investments. Comparing the losses experienced by children b and c, the difference is a matter of exposure rather than timing. Both children are exposed in the second period, however c is worse off if her parents are borrowing constrained in period 3.

These differences in timing and exposure provide a natural framework to think about the effects of parental unemployment on children and compare across children exposed at different times for different durations. For example, comparing the losses experienced by child a with c,

$$
\begin{aligned}
& \ln(Y(1,0)) - \ln(Y(0,0)) - (\ln(Y(1,1)) - \ln(Y(0,0))) \\
= \;& \left(\eta_2 - \eta_1 - \eta_2 \frac{\eta_1}{1-\eta_2-\frac{1}{\theta}}\right)\frac{i_{a2}-i_{a'2}}{i_{a'2}} + \left(\eta_1 - \eta_2 \frac{\eta_1}{1-\eta_2-\frac{1}{\theta}}\right)\left(\frac{i_{a2}-i_{a'2}}{i_{a'2}} - \frac{i_{c1}-i_{c'1}}{i_{c'1}}\right) \\
& - \frac{i_{b2}-i_{c2}}{i_{c2}} \frac{\gamma_2 i_{c2}^{\theta}}{\gamma_1 i_{c1}^{\theta} + \gamma_2 i_{c2}^{\theta}} \\
= \;& \underbrace{\left(\eta_2 - \eta_1 - \eta_2 \frac{\eta_1}{1-\eta_2-\frac{1}{\theta}}\right)\Delta\% i_{a2}}_{Timing} - \underbrace{\frac{\gamma_2 i_{c2}^{\theta}}{\gamma_1 i_{c1}^{\theta} + \gamma_2 i_{c2}^{\theta}} \frac{i_{b2}-i_{c2}}{i_{c2}}}_{Exposure}
\end{aligned}
$$

where the second equality follows from assuming $\Delta\% i_{a2} = \Delta\% i_{c1}$. Empirically, the scale of initial income losses from unemployment are generally independent of the child's age conditional on the year of exposure[45]. As such, one interpretation of this simplifying assumption is to consider a and c children who were exposed in the same year, but a is a year older than c.

A2 Identification

As in Section 2, let D_1 and D_2 be the parent's realized employment status in periods 1 and 2, were $D_a = 0$ describes employment and $D_a = 1$ describes unemployment in period a. Using Torgovitsky's notation, let $\{U_a(0), U_a(1)\}$ denote the potential outcomes for employment in D_a when $D_{a-1} = 0$ and $D_{a-1} = 1$ respectively. In particular,

$$
D_a = (1 - D_{a-1})U_a(0) + D_{a-1}U_a(1)
$$

Given an initial condition D_0 and $\{U_a(0), U_a(1)\}_{a=1,2}$ we can full describe a parent's employment history over the two periods. My object of interest is the child's income attainment at ages 27 to 31, Y. Let $Y(d_1, d_2)$ be the income attainment of a child when parental employment history is (d_1, d_2). Realized income attainment is given by,

$$
Y = \sum_{d_1 \in \{0,1\}} \sum_{d_2 \in \{0,1\}} Y(d_1, d_2)\mathbf{1}\{D_1 = d_1\}\mathbf{1}\{D_2 = d_2\}
$$

I estimate the effect of first experiencing parental unemployment at age 1 in the data. This amounts matching observations treated and control units on observables $x \in \mathbf{X}$ and then estimating in my matched sample, $\mathbb{E}(Y|D_1 = 1) - \mathbb{E}(Y|D_1 = 0)$

Under two assumptions, **A1.1** and **A1.2**, I can express $\mathbb{E}(Y|D_1 = 1) - \mathbb{E}(Y|D_1 = 0)$ as a weighted

[45] This has been noted empirically by Hilger (2016), and has been verified to be the case in my data as well.

average of four treatment effects of interest in the population,

$$Y(1,1) - Y(0,0)$$
$$Y(1,1) - Y(0,1)$$
$$Y(1,0) - Y(0,0)$$
$$Y(1,0) - Y(0,1)$$

A1 Common Support For every child who is first treated at age $a \in \{1, 2\}$ with observables $x \in \mathbf{X}$, there is a non-zero probability of not being treated.

$$\textbf{A1.1} \qquad Pr(D_1 = 1|X = x) \in [0, 1) \ \ \forall x \in \mathbf{X}$$
$$\textbf{A1.2} \qquad Pr(D_2 = 1|X = x, D_1 = 0) \in [0, 1) \ \ \forall x \in \mathbf{X}$$

A2 Conditional Independence of Potential Outcomes First experiencing treatment at age $a \in \{1, 2\}$ is as good as randomly assigned conditional on observables X.

$$\textbf{A2.1} \qquad D_1|X = x \perp\!\!\!\perp \{Y(0,1), Y(0,0), U_2(0), U_2(1)\} \ \ \forall x \in \mathbf{X}$$
$$\textbf{A2.2} \qquad D_2|X = x, D_1 = 0 \perp\!\!\!\perp \{Y(0,0)\} \ \ \forall x \in \mathbf{X}$$

As with any identification proof using a matching estimator, I first establish that I can estimate a weighted average of the four possible treatment effects in the population under **A2.1** at a given point $x \in \mathbf{X}$. Under **A1.1**, it follows that I can estimate these averages at every point of support where $Pr(D_1 = 1|X = x) \in (0, 1)$. As in Section 2, it is helpful to sort parents into four categories based on their potential outcomes $\{U_2(0), U_2(1)\}$,

	$U_2(0) = 0$	$U_2(0) = 1$
$U_2(1) = 0$	Always employed (E)	Negative state dependence (NS)
$U_2(1) = 1$	Positive state dependence (PS)	Always unemployed (U)

Notice that each of these groups is associated with one of the four possible treatment effects of interest for a child first exposed at age 1. In particular:

PS: Children whose parents experience positive state dependence and are assigned to $D_1 = 1$ will realize $Y = Y(1,1)$ due positive state dependence (i.e. $U_2(1) = 1$). If these children had counterfactually experienced $D_1 = 0$, positive state dependence would have had them realize $Y = Y(0,0)$.

U: Children whose parents would have always been unemployed in period 2 (i.e. $U_2(0) = U_2(1) = 1$) and are assigned to $D_1 = 1$ will realize $Y = Y(1,1)$. If they counterfactually experienced $D_1 = 0$, then they would have realized $Y = Y(0,1)$.

E: Children whose parents would have always been employed in period 2 (i.e. $U_2(0) = U_2(1) = 0$) and are assigned to $D_1 = 1$ will realize $Y = Y(1,0)$. If they counterfactually experienced $D_1 = 0$, then they would have realized $Y = Y(0,0)$.

NS: Children whose parents experience negative state dependence and are assigned to $D_1 = 1$ will realize $Y = Y(1,0)$ due negative state dependence (i.e. $U_2(1) = 0$). If these children had counterfactually experienced $D_1 = 0$, negative state dependence would have had them realize $Y = Y(0,1)$.

In the data $\forall x \in \mathbf{X} : Pr(D_1 = 1|X = x) \in (0,1)$ we have,

$$
\begin{aligned}
\mathbb{E}(Y|D_1 = 1, X = x) \quad = \quad & Pr\left[PS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,1)|D_1 = 1, PS, X = x\right] \\
& + Pr\left[U|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,1)|D_1 = 1, U, X = x\right] \\
& + Pr\left[E|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,0)|D_1 = 1, E, X = x\right] \\
& + Pr\left[NS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,0)|D_1 = 1, NS, X = x\right]
\end{aligned}
$$

where $Pr\left[G|D_1 = 0, X = x\right] = Pr\left[U_0(1) = u, U_2(1) = u'|D_1 = 0, X = x\right]$ for each group of parents $G \in \{PS, U, E, NS\}$ which is defined by $U_0(1) = u \wedge U_0(1) = u'$ for $u, u' \in \{0,1\}$. Under Assumption **A2.1** realized treatment status in period 1 is independent of potential outcomes conditional on $X = x$, so

$$
\begin{aligned}
\mathbb{E}(Y|D_1 = 0, X = x) \quad = \quad & Pr\left[PS|D_1 = 0, X\right]\mathbb{E}\left[Y(0,0)|D_1 = 0, PS, X = x\right] \\
& + Pr\left[U|D_1 = 0, X = x\right]\mathbb{E}\left[Y(0,1)|D_1 = 0, U, X = x\right] \\
& + Pr\left[E|D_1 = 0, X = x\right]\mathbb{E}\left[Y(0,0)|D_1 = 0, E, X = x\right] \\
& + Pr\left[NS|D_1 = 0, X = x\right]\mathbb{E}\left[Y(0,1)|D_1 = 0, NS, X = x\right] \\
= \quad & Pr\left[PS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(0,0)|D_1 = 1, PS, X = x\right] \\
& + Pr\left[U|D_1 = 1, X = x\right]\mathbb{E}\left[Y(0,1)|D_1 = 1, U, X = x\right] \\
& + Pr\left[E|D_1 = 1, X = x\right]\mathbb{E}\left[Y(0,0)|D_1 = 1, E, X = x\right] \\
& + Pr\left[NS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(0,1)|D_1 = 1, NS, X = x\right]
\end{aligned}
$$

At each $x \in X$, $\mathbb{E}(Y|D_1 = 1, X = x) - \mathbb{E}(Y|D_1 = 0, X = x)$ estimates a weighted average of the treatment effects for each group $G \in \{PS, U, E, NS\}$. While the weights are not identified in the the data (we never jointly observe $U_2(0)$ and $U_2(1)$) , they correspond to the likelihood of observing a parent who is a member of group G according to her potential outcomes at each point $D_1 = 1, X = x$. Specifically,

$$\begin{aligned}
\mathbb{E}(Y|D_1 = 1, X = x) - \mathbb{E}(Y|D_1 = 0, X = x) \quad = \quad & Pr\left[PS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,1) - Y(0,0)|D_1 = 1, PS, X = x\right] \\
& + Pr\left[U|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,1) - Y(0,1)|D_1 = 1, U, X = x\right] \\
& + Pr\left[E|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,0) - Y(0,0)|D_1 = 1, E, X = x\right] \\
& + Pr\left[NS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,0) - Y(0,1)|D_1 = 1, NS, X = x\right]
\end{aligned}$$

Under **A1.1**, we can integrate over all $x \in \mathbf{X}$ such that the probability of treatment is non-zero, resulting in my estimated effect,

$$\begin{aligned}
\mathbb{E}(Y|D_1 = 1) - \mathbb{E}(Y|D_1 = 0) \quad = \quad & \mathbb{E}_X\left[Pr\left[PS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,1) - Y(0,0)|D_1 = 1, PS, X = x\right]\right] \\
& + \mathbb{E}_X\left[Pr\left[U|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,1) - Y(0,1)|D_1 = 1, U, X = x\right]\right] \\
& + \mathbb{E}_X\left[Pr\left[E|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,0) - Y(0,0)|D_1 = 1, E, X = x\right]\right] \\
& + \mathbb{E}_X\left[Pr\left[NS|D_1 = 1, X = x\right]\mathbb{E}\left[Y(1,0) - Y(0,1)|D_1 = 1, NS, X = x\right]\right]
\end{aligned}$$

The identified effect for children first exposed at age 2 is considerably simpler, as age 2 is the terminal period of childhood in Section 2. As a consequence, there is no need to consider the parent's potential employment outcomes in this case. Following the same logic, under **A1.2** and **A2.2** it can be shown,

$$\mathbb{E}(Y|D_1 = 0, D_2 = 1) - \mathbb{E}(Y|D_1 = 0, D_2 = 0) \quad = \quad \mathbb{E}\left[Y(0,1) - Y(0,0)|D_1 = 1\right]$$

A3 Matching Appendix

Assumptions **A1** and **A2** do not readily apply to the data. In the case of continuous matching variables, the probability of two observations realizing the exact same value of a given variable is infinitesimal. When evaluating whether one untreated child is a valid counterfactual for another, in the sense of satisfying **A2**, researchers to deploy all signal variables at their disposal to generate children with the closest $Pr(D_i(a)|X = x_{j(i)})$. Of course, researchers never know how informative each matching variable is *ex ante*, therefore have incentives to hedge by including as much information on treatment status as possible. Doing so incurs the curse of dimensionality, and results in observations which are further and further apart as more matching variables are included. A successful application of any matching estimator needs to balance these two countervailing needs.

In this Section, I compare my matching strategy with caliper matching and coarsened exact matching in detail. The main advantage of proximity matching is it focuses on signal variables to create matches. The random forest performs data-driven dimension reduction, rather than having the researcher select which dimensions are relevant for matching. With the number potentially relevant of matching variables in administrative data, this is a large advantage. Moreover, my main specification in Section 5 cannot accommodate individual-level fixed effects. In other applications (e.g. Stepner, 2019), these fixed effects absorb time invariant heterogeneity across individuals and therefore their estimates only require balance on the trends within individuals to satisfy **A2**. In the absence of these fixed effects, I rely solely on the quality of my matches to ensure my estimates satisfy **A2**. As the random forest makes more efficient use of all information in the data than conventional methods, it is best suited to satisfy **A2**[46].

The empirical literature employing matching estimators to study the effects of layoffs has settled on deploying either caliper matching (Stepner, 2019) or coarsened exact matching Morthorst et al. (2021). In the case of the former, researchers select a vector $x_{j(i)} \in X \subset \mathbb{R}^k$ of observables to match individuals on, then calculate the pairwise distance between every treated observation and every possible control observation[47]. Unique matches are constructed by finding the closest control observation each treated observation, subject to a maximum caliper distance. Constructing matches in this fashion is agnostic of the treatment propensity function itself. In some sense, caliper matching effectively assumes that the treatment propensity function is locally smooth, so nearby observations will have very similar probabilities of treatment and be more likely to satisfy conditional independence of potential outcomes[48]. As a result, caliper matching avoids the misspecification error which can come along with propensity matching, at the cost of having to select all relevant matching dimensions while a caliper high enough to create enough matches for inference and low enough to ensure matches satisfy **A2**.

Coarsened exact matching strategies employ a similar logic, using a spiked distance measure rather than a smooth one. To create coarsened exact matches, researchers discretize any continuous matching variables, create bins for all combinations of the discrete matching variables, and then match randomly within these bins. For example, one application could sort all workers into bins defined by income deciles, a discretized continuous variable, and educational attainment, a naturally discrete variable, then match randomly within these bins. As with caliper matching, coarsened exact matching satisfies **A2** so long as the potential outcomes of individuals in these bins are not related to their realized treatment status. Again, if we think of a Roy model were individuals select into treatment according to their private knowledge of their potential outcomes, it can be informative to think of what coarsened exact matching assumes in terms of the treatment propensity function. While coarsened exact matching also avoids imposing a specification on the propensity function, it requires researchers to make different *ex ante* decisions when selecting where to cut continuous variables rather a maximum caliper. One key advantage of this approach, however, is that it significantly reduces the

[46] In Section A5, I present my primary test of **A2** when I match using caliper matching. When I include individual-level fixed effects, my results largely satisfy **A2**, suggesting caliper matching is sufficient for researchers whose identification relies on matching treatment and control trends within an individual. In the absence of these fixed effects, this is not the case.

[47] With infinite data, the choice of norm to deploy $\mathbb{R}^k$ should not affect results. In practice, researchers typically use Mahalanobis distance (Stepner, 2019) as their norm of choice, as it adjusts for the covariance between the different dimensions of $x_{j(i)}$.

[48] This is true, for example, if treatment propensity contains some information about the selection decisions made by individuals according to their potential outcomes.

computational complexity of the problem by sorting the data into thousands of bins rather than computing $n_{control} \times n_{treated}$ distances[49]. In practice we expect the matrix of pairwise distances employed in caliper matching to be relatively sparse, and as a result many papers will borrow from the coarsened exact match logic by taking a "blocking" and only consider matching observations with the same realization of a discrete variable (e.g. from the same income decile in Stepner, 2019).

In my preferred approach, I employ a procedure related to both caliper matching and coarsened exact matching. I use a random forest to predict the treatment status of fathers, and match observations using the forest's proximity matrix as my measure of how close two observations are to one another. Proximity matching has been employed in other fields to recover causal effects (e.g. Zhao et al., 2016), and fits with a larger econometric literature using random forests for causal inference and estimation of heterogenous treatment effects (Athey and Imbens, 2016; Wager and Athey, 2018; Athey et al., 2018; Friedberg et al., 2021). The primary advantage of proximity matching is that random forests implicitly perform dimension reduction, and the resulting matches are created using a distance measure which focusses on signal dimensions in the space of matching variables.

The proximity matrix arises from the many decision trees which make up a given random forest. Each tree uses a randomly selected subsample of the data to categorize observations into treatment or control by sequentially partitioning the space of matching variables into hyperrectangles to predict treatment status[50]. A forest then generates predictions of treatment by sorting a observation i into a hyperrectangle for each tree given its realization of observables $x_{j(i)}$, and predicting its treatment status based on whether $x_{j(i)}$ falls into hyperrectangles which only contain treated or control observations the majority of the time. In this way, random forests can be understood as a class of nearest-neighbour estimators (Lin and Jeon, 2006; Biau and Devroye, 2010; Friedberg et al., 2021) and two observations i and i' with observables $x_{j(i)}$ and $x_{j(i')}$ will receive similar predictions from a forest the more often their observables fall in the same hyperrectangle.

The proximity matrix summarizes how often two observations i and i' fall into the same hyperrectangle in a given random forest[51]. In particular, it is a $n \times n$ matrix which details the fraction of the trees where i and i' are sorted into the same hyperrectangle. Two observations i and i' with higher proximity will also be more similar their predicted treatment status, as a result of being close along dimensions in X which are predictive of treatment status. I employ this measure to uniquely match every treated observation i to the control observation i' highest in proximity. Similar to caliper matching, I do not place any restrictions on the treatment propensity function nor do I restrict to matching i to a counterfactual i' which has the same predicted treatment status. Similar to coarsened exact matching, proximity matching relies on a spiked distance

[49]These costs can be very high with administrative data. For example, even a relatively small administrative dataset of 150,000 treated observations and 150,000 potential control observations will require at least 22.5 GB of RAM for the matrix of all pairwise comparisons between treated and control observations – approaching the RAM limits of most personal computers. Calculating the distance between two observations to several decimal places would increase that RAM requirement well beyond the single bit to store a 0 or 1 which constitutes this minimum.

[50]For an introduction to random forests geared towards economists, see Athey and Imbens (2019). Biau and Scornet (2015) provides an excellent introduction discussing the underlying algorithms and statistical properties of random forests.

[51]Olson and Wyner (2018) explicitly relate the proximity matrix to nearest neighbour estimators by proposing the "proximity kernel" which casts random forest predictions as resulting from a kernel estimator.

measure for each tree and requires all specified dimensions fall within a given range for two observations to be considered close. Unlike caliper matching, matching variables are not selected by the researcher *ex ante*, nor are the boundaries of the hyperrectangles as they are in coarsened exact matching. Proximity matching takes a data-driven approach to selecting which matching variables contain signal variation on treatment status, and should be used when calculating how similar two observations are. Similarly, each tree selects which partitions of the space of these observables contain the most information on treatment status, and are therefore best employed when predicting treatment status.

Figure A1 provides a stylized example of a proximity matrix. Suppose that our forest is made up of the three trees presented in Panel A, and that observations 1 through 8 happened to be "out of bag" (not included in the subsample) when each of these trees were grown. Among the matching variables available to predict treatment status at each split, all trees used the x_1 and x_2 to partition the space of observables. Observations 1 to 8 can then be sorted into their respective rectangles given their realizations of x_1 and x_2 as presented in Panel A. Calculating the fraction of trees where a pair of observations appear in the same rectangle then yields the proximity between the pair, presented in the matrix in Panel B. For example, observations 4 and 6 both fall together in the top right rectangles in Trees 1 and 3, but they end up in sorted into separate rectangles in Tree 2. Therefore, the proximity between 4 and 6 is $\frac{2}{3}$.

It is important to note that the proximity matrix presented in Figure A1 is generated only using out of bag observations. A concern which frequently arises in the econometric literature using random forests for causal inference is whether estimates are model dependent (Athey and Imbens, 2016). In particular, if the proximity between i and i' is calculated using trees where either observation was included in the subsample used to grow the tree, then matching is dependent on the observations themselves and confidence intervals may not provide appropriate coverage. Following Athey and Imbens (2016), I ensure that my proximity matching strategy is "honest" and all matches are independent of the observations themselves. In particular, I calculate the proximity for every pair of observations i and i' only using the trees where the are jointly out of bag. As shown in Athey and Imbens (2016), using predictions generated when observations are out of bag guarantees valid inference and that standard asymptotics apply to a given matched sample.

Finally, notice that Figure A1 contains no information on which rectangles predict observations will realize $D_a = 0$ or $D_a = 1$. While the hyperrectangles created by a given tree are created to predict treatment status, the proximity matrix they create only measures how close two observations are in signal dimensions. This distinction is important, as it differentiates proximity matching from propensity score matching. Propensity score matching procedures estimate the treatment propensity function, potentially using a random forest to do so, and and then match observables with similar predicted treatment probabilities. While my proximity matching procedure also estimates a treatment propensity function, I do not use the predicted treatment status or predicted probability of treatment to create matches. This avoids one of the main disadvantages of propensity matching — it requires all of the information on the similarities between observations to be collapsed into a single prediction of treatment probability, resulting in information loss (King et al., 2011;

King and Nielsen, 2019). For example, two observations could have completely different observables along matching variables, but still receive the same predicted probability of treatment. By ignoring how (dis)similar these observations are along observables in the matching procedure, propensity score matching could result in these two observables being matched. Much like caliper and coarsened exact matching, proximity matching avoids this issue by focussing on how close two observations are in terms of matching variables rather than their predicted treatment propensity.

A4 Random Forest Calibration for Proximity Matching

Though observations with high proximity will not necessarily receive the same predicted treatment status from a random forest, the quality of the predictions is intimately related to the quality of measured proximity. The predictions of random forests are affected by the size of subsample available to each tree n_s, the number of predictors considered at each node to partition the data *mtry*, and the maximum number of unique observations allowed in a terminal node of the forest *nodesize*. In this Section, I address how these parameters affect my results.

Higher values of n_s for a forest containing a fixed number of trees will make it less likely that any two observations are jointly out of bag, and make proximity estimates reliant on less trees. For example, the probability of two observations being jointly out of bag in a given tree is $\prod_{i=1}^{n_s}\left(1-\frac{1}{n-i}\right)^2$ for a forest estimated using n observations when sampling without replacement. As $n_s \to n$, this probability approaches zero. Larger values of n_s will also lead to more terminal nodes for a fixed *nodesize*, more hyperrectangles in each for the out of bag observations to be sorted into, and therefore lower proximity overall. The same logic applies when using lower values of *nodesize* for a fixed subsample size n_s. Lower *nodesize* will create more leaves, making each leafs more granular, and lower proximity as a result.

Unlike n_s and *nodesize* which govern how many leafs are created by a tree, $m_t ry$ affects the quality of predictors chosen to form each leaf. In particular, higher values of *mtry* will make it more likely that a signal variable is available to be selected by a tree at each split, resulting in trees which focus on signal variation more than they would for lower values of *mtry*. Of course, at higher and higher values of *mtry* there will be be more homogeneity among the trees in the forest, which should encourage more extreme values of proximity. To see this, suppose that all matching variables are available at each split and subsamples drawn from the data are relatively homogenous. Each tree in the forest will choose similar partitions of the data as a result of having the same variables at its disposal, and therefore out of bag observations sorted into one hyperrectangle together are more likely to be sorted into all other hyperrectangles together. Biau and Scornet (2015) provides further discussion of how these hyperparameters affect the quality of a random forest's predictions.

In the results presented in this paper, I employ the default algorithm for classification trees with *mtry* set to the square root of the number of matching variables and *nodsize* is set to 1. To guarantee all observations are jointly out of bag in many trees, I set n_s equal to half of the the total observations available to the forest and grow 500 trees for each iteration. As discussed in Section A3, one key advantage of proximity matching over caliper or coarsened exact matching is that the matching selected in a data driven fashion rather than depending on a researcher's decisions. Tuning these hyperparameters is one such way researchers could affect the resulting matching, and as a result I opt not to do so in the results I present in the paper.

An alternative strategy would be to tune the forest focussing on a hyperparameter which reflects the quality of the forest's prediction. Two commonly targeted hyperparameters in random forest applications are sensitivity (true positive rate among predicted positives) and specificity (true negative rate among predicted negatives)[52]. In previous versions of this paper, I have tuned the forest to maximize the minimum of sensitivity and specificity, with few qualitative differences in the results.

A5 Alternative Matching Strategies

In this Section, I consider how different matching strategies would affect my results. I horse race proximity matching against caliper matching in my data. I show that both procedures create more comparable match pairs when researchers are able to use individual-level fixed effects to absorb cross-sectional heterogeneity in the data. When these fixed effects are not included, there is a considerable reduction in balance in the caliper matched sample, while the equivalent tests using the proximity matched sample yield comparable results in both cases. For the proximity matched sample, I reject balance tests with the expected Type I error rate regardless of whether fixed effects are included or not (rejection rate of 4.1% and 5% respectively at the 5% level). For the caliper matched sample, I over reject balance tests compared to the Type I error rate when fixed effects are and are not included (rejection rates of 11.4% and 37.7% respectively at the 5% level).

To create a comparable caliper matched sample, I replicate the procedure used in Stepner (2019) in my data. To do so, I first remove all chronic claimants using the same procedure as described in Section 4.1. I then sort each same of possible treated and control children of age a in year t into income deciles according to their father's taxable income, and evaluate all pairwise distances between the observables of a treated father, $x_{j(i)}$, and possible control father, $x_{j(i')}$. In particular, I evaluate the Mahalanobis distance[53] between all treated and control fathers and then match each treated father to closest the control father without replacement. Following Stepner (2019), I throw out all matched observations with a Mahalanobis distance beyond a maixmum (caliper) distance of 1. As discussed in Section A3, the curse of dimensionality precludes me from including all of the same matching variables in the caliper matching as in the random forest[54].

[52] If we have the null hypothesis that no fathers are unemployed, one familiar way to interpret these parameters resepctively is 1-Type I errors in prediction and 1-Type II errors in prediction.

[53] Given by $d(x_{j(i)}, x_{j(i')}) = \sqrt{(x_{j(i)} - x_{j(i')})' S^{-1} (x_{j(i)} - x_{j(i')})}$ where S is the sample variance-covariance matrix of all dimensions of x for the relevant block of the data.

[54] In particular, including more information which we think has little value in predicting treatment status, e.g. professional income, will result in fewer matches forming within a fixed caliper width. This is a known disadvantage of caliper matching, and requires some tuning on the part of the researcher.

Instead, I follow Stepner (2019) and match on father's year of birth, province of residence in $t-3$, marital status in $t-3$, number of children, years without claiming UI in $t-3$, as well as the father and household's net income percentile in $t-3$, taxable income in $t-3$, and labour income $t-3$.

In Section 4.3, I argue that I can test for balance between treated and control children in my matched sample by focussing at their parents' observables, $\theta_{j(i)}$, prior to matching in year $t-3$. In equation (5), this means running a regression of the form,

$$\theta_{j(i)t} = \sum_{t} \mu_t + \delta_t D_i(a) + \epsilon_{j(i)t}$$

and testing $H_0 : \delta_t = 0 \ \ \forall t < \tau - 3$. As I estimate 20 treatment effects, one for every age of exposure, I perform this test separately for my matched samples of children treated at age $a \in \{2, 3, ..., 21\}$.

In Tables 4 and A5, I present p-values which result from testing these hypotheses in the proximity and caliper matched samples. In Table 4, I reject 11 of 220 null hypotheses at the 5% level, in line with expected Type I error rates. In Table A5, I reject the same null hypotheses in the caliper matched sample in 83 of 220 tests at the 5% level. Among the rejections, it is clear that the fathers' spouses are extremely different between treatment and control observations in these samples, consistent with a story of marriage providing income insurance and affecting risk taking (Chiappori et al., 2018). Of course, the spouse's information was not made available to the caliper match beyond household income, so one could argue that this is an example of a need to use more matching variables rather than an issue with caliper matching per se. Focussing in on the fathers, however, I reject 11 of 120 null hypotheses at the 5% level — again well above the expected Type I error rate, despite the father's information being the focus of the caliper match.

To understand why caliper matching under performs relative to proximity matching, I consider including individual-level fixed effects in my test specification to absorb time invariant, cross-sectional differences between the matched treatment and control group. In particular, I run regressions of the form,

$$\theta_{j(i)t} = \chi_{j(i)} + \sum_{t \neq \tau - 3} \tilde{\mu}_t + \tilde{\delta}_t D_i(a) + \epsilon_{j(i)t}$$

where $\chi_{j(i)}$ is a parent-level fixed effect, and test $H_0 : \tilde{\delta}_t = 0 \ \forall t < \tau - 3$.

Tables A4 and A6 present the results when using the proximity matched sample and the caliper matched sample respectively. Using the proximity matched sample, I reject 9 of the 220 tests at the 5% level, again well in line with expected Type I error rates. The inclusion of fixed effects in Table A6 considerably improves balance when compared with A5. I reject 25 of 220 null hypotheses tested at the 5% level, a noticeable improvement, but a rejection rate which is more than double the expected Type I rate. This suggests that many of the differences between treatment and control groups in Table A6 arose from cross-sectional differences in the matched sample, rather than pre-trends within a child's household. Focussing on the father's information, I reject 12 of 120 hypotheses tested at the 5% level. This rejection rate is largely unchanged when compared with Table A6, and still considerably higher than the expected Type I error rate.

Overall, it is apparent that the proximity matching outperforms a caliper matching in this application. Proximity matching delivers a more balanced matched sample, in part because the random forest's dimension reduction means that a proximity matching can accommodate more observable characteristics than caliper matching (or coarsened exact matching). There is considerably more balance between treatment and control in the caliper matching application when allowing for individual fixed effects. This suggests that the proximity matching may be better than caliper matching at matching individuals to achieve balance in unobserved, time invariant heterogeneity, consistent with the data-driven match definition in the random forest. As my main specification in equation (4) does not allow for individual fixed effects, a proximity matching is particularly well-suited to my setting.

A6 Treatment Effect Decomposition

To parse out the importance of these mediators, I employ a decomposition similar to Heckman et al. (2013). I use observables on fathers of treated children experienced at age a in year t, $\theta_{j(i)t}$, and the same information for matched counterfactuals, $\theta_{j(i')t}$, to understand how parental unemployment affects children. This approach allows for a given estimated effect Δ_a to be expressed in terms of its intermediate effects on input k, such as income lost $\theta^k_{j(i)} - \theta^k_{j(i')}$ at age a, and a residual difference in means driven by unobserved inputs to a child's income attainment, Y_i. Suppose that for each age of exposure to parental unemployment, all inputs are related to a child's income attainment according to an empirical model,

$$Y_i = \sum_k (\psi^k_{a,0}(1 - D_i(a)) + \psi^k_{a,1} D_i(a))\theta^k_{j(i)} + \zeta_i$$

where $\psi^k_{a,d}$ is the average return on input θ^k for children in treated at age a or their matched controls[55].

Following Heckman et al. (2013), suppose that we can partition the $k \in K$ into a set of observed inputs K_o and a set of unobserved inputs K_u. For any input $k \in K$, observable or unobservable, we can consider an empirical model of the form,

$$\theta^k_{j(i)} = \varphi_a + \sum_k \phi^k_a D_i(a) + \nu_j(i)$$

where ϕ^k_a is the expected difference in inputs $\theta^k_{j(i)}$ received by treated and control children. In particular, $\phi^k_a = \mathbb{E}[\theta^k_{j(i)}(1,a) - \theta^k_{j(i)}(0,a)]$, where $\theta^k_{j(i)}(0,a)$ and $\theta^k_{j(i)}(1,a)$ are respectively the inputs she would have experienced if she did not and did experience parental at age a.

In this setting, we can decompose an estimated effect $\hat{\Delta}_a$ into a portion explained by observed inputs $k \in K_o$ and a residual, driven by the unobserved inputs to a child's income attainment,

[55] We should expect that this return varies by the age of first exposure to parental unemployment, as Appendix A2 establishes that the nature of treatment varies depending on the year of exposure. Similarly, the composition of the matched control group varies by the year of exposure, so we should expect that the average return they face on input θ^k might as well.

$$\underbrace{\hat{\Delta}_a}_{Estimated\ Effect} = \underbrace{\sum_{k \in K_o} \hat{\psi}^k_{a,0} \hat{\phi}^k_a}_{Explained} + \underbrace{\sum_{k \in K_u} \hat{\psi}^k_{a,0} \hat{\phi}^k_a}_{Residual}$$

This decomposition requires three assumptions to have a causal interpretation. The first is a matter of interpretation. I assume that parental unemployment does not have a direct effect on children. Consequently, the unexplained residual is entirely explained by parental unemployment's effects in unobserved inputs (e.g. mental health in the household) rather than a direct effect on children. Other applications such as Heckman et al. (2013), it is natural to include a direct of treatment, and the residual term ought to be interpreted as the sum of direct effects of treatment and the indirect effects of treatment on unobserved inputs.

The second assumption is the most tenuous, and requires that $cov(\theta^k_{j(i)}, \theta^{k'}_{j(i)}) = 0$ for $k \in K_o$ and $k \in K_u$. Under this assumption, I can obtain consistent estimates of $\psi^k_{a,0}$ in equation (7) as $cov(\theta^k_{j(i)}, \theta^{k'}_{j(i)}) = 0$ implies there is no omitted variable bias. In the data, I implement equation (7) with match pair fixed effects to remove some of these potential omitted variables. As assumption **A2** states treatment is conditionally independent of potential outcomes, these fixed effects will absorb any cross sectional differences between children prior to treatment. Nonetheless, it is still reasonable to expect that an unobserved input to the child's income attainment, such as a parent's mental health, and observed inputs, such as income reductions caused by unemployment, are correlated within match pairs. As such, my mediation analysis is best understood as descriptive evidence rather than causal.

Finally, to arrive at this decomposition I must assume that $\psi^k_{a,0} = \psi^k_{a,1} \equiv \psi^k_a$. Unlike the preceding assumptions, this is testable in the data by evaluating whether $H_0 : \psi^k_{a,0} = \psi^k_{a,1}$. Heckman et al. (2013) show that this test is still valid even when the second assumption fails, and is equivalent to assuming that any omitted variables $k' \in K_u$ are equally related to the correlated with the input $k \in K_o$ for treated and control children. I test this assumption in the data using my main specification in equation (7), and fail to reject the null hypothesis $\psi^k_{a,0} = \psi^k_{a,1}$ at conventional levels for any age of exposure and input.

A7 Appendix: Figures and Tables

Table A1: Match-Year Summary Statistics

	Received UI Benefits		**Possible to Match**		**Matched Sample**	
	Yes	No	Treated	Control	Treated	Control
	(1)	(2)	(3)	(4)	(5)	(6)
Father						
Total income	31400	43500	33400	43600	40700	40500
	(17)	(14)	(27)	(15)	(144)	(149)
Labour income	30700	42200	33000	42600	42000	41500
	(15)	(12)	(24)	(13)	(121)	(116)
Father's Spouse						
Total income	13600	17100	13800	16800	18500	18800
	(163)	(11)	(382)	(12)	(113)	(106)
Labour income	14500	17300	15000	17200	20200	20700
	(14)	(5)	(22)	(6)	(106)	(103)
Household						
Total income	39400	54900	41700	54900	54100	54000
	(101)	(17)	(241)	(18)	(192)	(191)
N	1,310,200	13,988,900	550,200	13,235,100	31,200	31,200

Notes: This table presents means and standard errors (in brackets) for a number of observables from match years 1984-1990 for treatment years 1987-1993. All income variables are in 1990 Canadian dollars. Columns (1) and (2) report statistics in the match year for all tax filing fathers in the IID who respectively did and did not receive UI benefits in the treatment year. Columns (3) and (4) report statistics in the match year for all treated and control observations who could possibly be matched respectively. Columns (5) and (6) report report statistics in the match year for all treated and control observations in my preferred matched sample respectively (that is, all observations matched with a proximity greater than 0.7). All estimates are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Table A2: Effects of Parental Unemployment on Child's Income Attainment - Two Year Age Bins

Age of Exposure	**Income Rank Age 27 to 31** (1)	**Income Rank Age 30 to 34** (2)
2-3	-3.746^{**}	
	(1.513)	
4-5	-1.546	-1.332
	(0.946)	(2.527)
6-7	-2.352^{***}	-4.429^{***}
	(0.886)	(1.412)
8-9	-2.983^{***}	-2.649^{***}
	(0.771)	(0.928)
10-11	-3.018^{***}	-2.200^{***}
	(0.779)	(0.800)
12-13	-2.955^{***}	-2.349^{***}
	(0.789)	(0.795)
14-15	-1.569^{**}	-1.812^{**}
	(0.737)	(0.719)
16-17	-2.499^{***}	-2.172^{**}
	(0.877)	(0.849)
18-19	-1.587	-1.179
	(1.086)	(1.074)
20-21	-0.314	-0.124
	(1.782)	(1.830)
N	37,500	28,700
Mean Control Outcome	53.9	53.4

Notes: This table presents point estimates of Δ_a from the regression presented in equation (4), controlling for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \ \forall a$). Standard errors presented in parentheses are clustered at the match-pair level. These estimates are also presented in Figure A2. All estimates are for 2 year age bins, with one exception:in Column 2, the first non-empty row only covers children aged 5 at exposure to parental job loss. *, **, and *** respectively indicate statistical significance at the 10%, 5%, and 1% levels. All estimates are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Table A3: Effects of Parental Unemployment on Child's Income Attainment - Age of Exposure

Age of Exposure	**Income Rank Age 27 to 31** (1)	**Income Rank Age 30 to 34** (2)
2	-4.954^{**}	
	(2.460)	
3	-3.116	
	(1.911)	
4	-2.073	
	(1.360)	
5	-1.152	-1.332
	(1.303)	(2.527)
6	-0.903	-2.282
	(1.324)	(2.379)
7	-3.424^{***}	-5.442^{***}
	(1.190)	(1.744)
8	-2.131^{*}	-1.825
	(1.102)	(1.402)
9	-3.787^{***}	-3.255^{***}
	(1.079)	(1.236)
10	-4.844^{***}	-3.656^{***}
	(1.092)	(1.177)
11	-1.014	-0.747
	(1.110)	(1.085)
12	-3.335^{***}	-2.173^{*}
	(1.169)	(1.208)
13	-2.641^{**}	-2.491^{**}
	(1.068)	(1.055)
14	-1.074	-1.583
	(1.033)	(1.009)
15	-2.029^{**}	-2.029^{**}
	(1.049)	(1.023)
16	-3.291^{***}	-2.459^{**}
	(1.160)	(1.125)
17	-1.508	-1.818
	(1.337)	(1.292)
18	-1.171	-0.925
	(1.405)	(1.416)
19	-2.268	-1.572
	(1.707)	(1.637)
20	-0.247	-0.493
	(2.186)	(2.229)
21	-0.441	0.547
	(3.079)	(3.190)
N	37,500	28,700
Mean Control Outcome	53.9	53.4

Notes: This table presents point estimates of Δ_a from the regression presented in equation (4), controlling for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \ \forall a$). Standard errors presented in parentheses are clustered at the match-pair level. These estimates are also presented in Figure A2. All estimates by age of exposure to parental job loss. *, **, and *** respectively indicate statistical significance at the 10%, 5%, and 1% levels. All estimates are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.

Table A4: Childhood Pretrends

	Father's Income			**Father's Spouse's Income**			**Household**		**Filed Taxes**		
Age	Taxable	Labour	Net of Tax	Taxable	Labour	Net of Tax	Married	Moved CSD	Father	Father's Spouse	Father's Spouse T4
2	0.8800	0.2475	0.9007	0.4497	0.2155	0.4157	0.9545	0.4421	0.8780	0.6078	0.3928
3	0.2736	0.1487	0.2743	0.7208	0.7889	0.7260	0.7096	0.2920	0.3289	0.1733	0.4492
4	0.2905	0.1156	0.2459	0.0261	0.1069	0.0229	0.6915	0.0025	0.1173	0.3314	0.8037
5	0.4659	0.6006	0.5153	0.3180	0.6743	0.3188	0.3755	0.6994	0.6736	0.6114	0.6889
6	0.5659	0.5930	0.6349	0.3484	0.8025	0.3458	0.7962	0.4429	0.6572	0.7166	0.5706
7	0.8732	0.3623	0.8657	0.5266	0.9350	0.5446	0.6601	0.7844	0.7175	0.8239	0.2931
8	0.5638	0.5740	0.5607	0.7100	0.4396	0.6828	0.7667	0.6632	0.7124	0.5253	0.8407
9	0.3471	0.0654	0.3746	0.8912	0.8437	0.8885	0.2396	0.1619	0.8095	0.4972	0.7283
10	0.0853	0.1465	0.0976	0.4263	0.5729	0.4218	0.3311	0.2364	0.6230	0.6052	0.5334
11	0.1472	0.1619	0.1078	0.8256	0.9239	0.8177	0.9657	0.1387	0.9607	0.6850	0.2157
12	0.4818	0.0731	0.5401	0.2729	0.5011	0.3170	0.8722	0.5695	0.8985	0.6631	0.2449
13	0.2125	0.0770	0.2074	0.4528	0.8686	0.3969	0.9541	0.2853	0.0583	0.9983	0.6362
14	0.7674	0.5056	0.7567	0.3465	0.4431	0.3464	0.3602	0.6596	0.8837	0.9688	0.9055
15	0.0305	0.2412	0.0421	0.3025	0.9766	0.3060	0.4189	0.3641	0.8906	0.9266	0.0293
16	0.2910	0.2390	0.1492	0.5708	0.4952	0.5684	0.5289	0.7374	0.5144	0.8949	0.8070
17	0.5677	0.1471	0.4675	0.7640	0.7752	0.8035	0.3931	0.1348	0.4514	0.8353	0.7713
18	0.0636	0.1744	0.0837	0.2361	0.6026	0.2440	0.9771	0.5936	0.6658	0.3405	0.9649
19	0.7235	0.5071	0.6667	0.9769	0.5638	0.9814	0.4803	0.0561	0.1012	0.8249	0.8734
20	0.9373	0.2748	0.8884	0.0872	0.0782	0.0821	0.1165	0.1808	0.9450	0.4793	0.3867
21	0.0374	0.0117	0.0431	0.3522	0.2999	0.3514	0.5142	0.2803	0.5442	0.7443	0.1833

Notes: These are the resulting p-values from testing the δ_t $\forall t < \tau - 3$ in equation (5) for a given outcome while also including child/parent-level fixed effects, with $\delta_{\tau-3}$ and $\mu_{\tau-3}$ both normalized to zero to permit the fixed effects. Each column presents a different outcome, and each row presents age of exposure to parental unemployment for the matched treatment and control sample. Standard errors are clustered at the level of child i. Labour income is measured on the T4 form, and excludes self-employment income as a result. The "Married" outcome includes married or co-habitating partners. Measures of whether father or father's spouse filed taxes are dummy variables for whether taxes records are available in a given year. Similarly, "Father's Spouse T4" is a dummy variable for whether the father's spouse received any labour income in a given tax year, measured on the T4 form.

Table A5: Childhood Pretrends - Caliper Matches

	Father's Income			**Father's Spouse's Income**			**Household**		**Filed Taxes**		
Age	Taxable	Labour	Net of Tax	Taxable	Labour	Net of Tax	Married	Moved CSD	Father	Father's Spouse	Father's Spouse T4
2	0.166	0.160	0.150	0.000	0.000	0.000	0.177	0.781	0.767	0.128	0.310
3	0.090	0.004	0.086	0.000	0.000	0.000	0.409	0.118	0.343	0.023	0.811
4	0.398	0.218	0.376	0.000	0.000	0.000	0.002	0.719	0.809	0.036	0.456
5	0.351	0.457	0.347	0.000	0.000	0.000	0.038	0.405	0.829	0.000	0.978
6	0.054	0.153	0.061	0.000	0.000	0.000	0.479	0.426	0.356	0.080	0.016
7	0.059	0.101	0.048	0.000	0.000	0.000	0.012	0.640	0.784	0.005	0.650
8	0.289	0.070	0.333	0.000	0.000	0.000	0.433	0.330	0.736	0.001	0.291
9	0.280	0.137	0.276	0.000	0.000	0.000	0.181	0.491	0.856	0.005	0.071
10	0.054	0.199	0.068	0.000	0.000	0.000	0.731	0.571	0.609	0.565	0.379
11	0.017	0.091	0.015	0.000	0.000	0.000	0.320	0.567	0.776	0.094	0.004
12	0.371	0.019	0.355	0.000	0.000	0.000	0.731	0.722	0.646	0.185	0.017
13	0.013	0.029	0.012	0.000	0.000	0.000	0.812	0.020	0.540	0.043	0.026
14	0.124	0.334	0.116	0.000	0.000	0.000	0.537	0.820	0.662	0.260	0.000
15	0.071	0.124	0.069	0.000	0.000	0.000	0.550	0.060	0.812	0.270	0.142
16	0.049	0.474	0.056	0.000	0.000	0.000	0.116	0.438	0.317	0.022	0.005
17	0.136	0.109	0.127	0.000	0.000	0.000	0.695	0.727	0.340	0.338	0.010
18	0.057	0.795	0.060	0.000	0.000	0.000	0.127	0.521	0.455	0.004	0.053
19	0.250	0.297	0.264	0.000	0.000	0.000	0.775	0.903	0.862	0.123	0.050
20	0.830	0.762	0.865	0.000	0.000	0.000	0.838	0.105	0.624	0.170	0.162
21	0.739	0.237	0.703	0.221	0.221	0.236	0.238	0.929	0.526	0.895	0.435

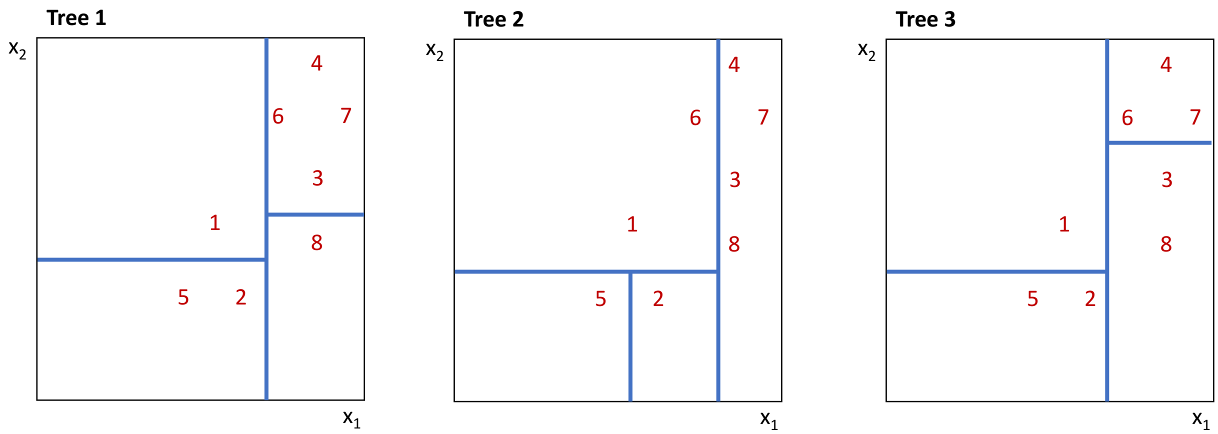
Notes: These are the resulting p-values from testing the $\delta_t \; \forall t < \tau - 3$ in equation (5) for a given outcome. Each column presents a different outcome, and each row presents age of exposure to parental unemployment for the matched treatment and control sample. Standard errors are clustered at the level of child i. Labour income is measured on the T4 form, and excludes self-employment income as a result. The "Married" outcome includes married or co-habitating partners. Measures of whether father or father's spouse filed taxes are dummy variables for whether taxes records are available in a given year. Similarly, "Father's Spouse T4" is a dummy variable for whether the father's spouse received any labour income in a given tax year, measured on the T4 form.

Table A6: Childhood Pretrends - Caliper Matches

	Father's Income			**Father's Spouse's Income**			**Household**		**Filed Taxes**		
Age	Taxable	Labour	Net of Tax	Taxable	Labour	Net of Tax	Married	Moved CSD	Father	Father's Spouse	Father's Spouse T4
2	0.297	0.341	0.291	0.325	0.191	0.315	0.112	0.416	0.767	0.362	0.175
3	0.032	0.006	0.036	0.183	0.089	0.180	0.302	0.279	0.397	0.616	0.887
4	0.282	0.154	0.288	0.526	0.674	0.545	0.001	0.914	0.811	0.316	0.438
5	0.325	0.019	0.307	0.278	0.270	0.308	0.021	0.311	0.883	0.567	0.956
6	0.490	0.058	0.570	0.026	0.075	0.030	0.361	0.590	0.285	0.486	0.068
7	0.077	0.051	0.097	0.397	0.168	0.386	0.002	0.488	0.713	0.184	0.462
8	0.369	0.410	0.430	0.048	0.044	0.038	0.156	0.061	0.824	0.765	0.211
9	0.266	0.026	0.292	0.321	0.531	0.301	0.153	0.634	0.937	0.025	0.009
10	0.110	0.314	0.169	0.336	0.148	0.376	0.875	0.910	0.515	0.545	0.548
11	0.156	0.020	0.173	0.454	0.203	0.445	0.355	0.643	0.710	0.521	0.003
12	0.208	0.157	0.214	0.004	0.060	0.004	0.890	0.871	0.786	0.790	0.034
13	0.413	0.009	0.449	0.321	0.041	0.349	0.763	0.008	0.522	0.380	0.007
14	0.038	0.694	0.049	0.748	0.271	0.741	0.623	0.425	0.633	0.346	0.000
15	0.106	0.088	0.139	0.252	0.414	0.214	0.488	0.151	0.820	0.371	0.845
16	0.080	0.011	0.105	0.534	0.354	0.528	0.070	0.341	0.335	0.220	0.140
17	0.617	0.008	0.600	0.597	0.384	0.612	0.695	0.804	0.302	0.441	0.311
18	0.207	0.148	0.207	0.101	0.265	0.137	0.222	0.783	0.629	0.131	0.169
19	0.217	0.545	0.204	0.294	0.384	0.299	0.766	0.771	0.778	0.051	0.246
20	0.868	0.631	0.886	0.183	0.312	0.165	0.832	0.255	0.562	0.274	0.337
21	0.830	0.198	0.836	0.470	0.234	0.493	0.234	0.897	0.526	0.886	0.562

Notes: These are the resulting p-values from testing the δ_t $\forall t < \tau - 3$ in equation (5) for a given outcome while also including child/parent-level fixed effects, with $\delta_{\tau-3}$ and $\mu_{\tau-3}$ both normalized to zero to permit the fixed effects. Each column presents a different outcome, and each row presents age of exposure to parental unemployment for the matched treatment and control sample. Standard errors are clustered at the level of child i. Labour income is measured on the T4 form, and excludes self-employment income as a result. The "Married" outcome includes married or co-habitating partners. Measures of whether father or father's spouse filed taxes are dummy variables for whether taxes records are available in a given year. Similarly, "Father's Spouse T4" is a dummy variable for whether the father's spouse received any labour income in a given tax year, measured on the T4 form.

Figure A1: Stylized Proximity Matrix

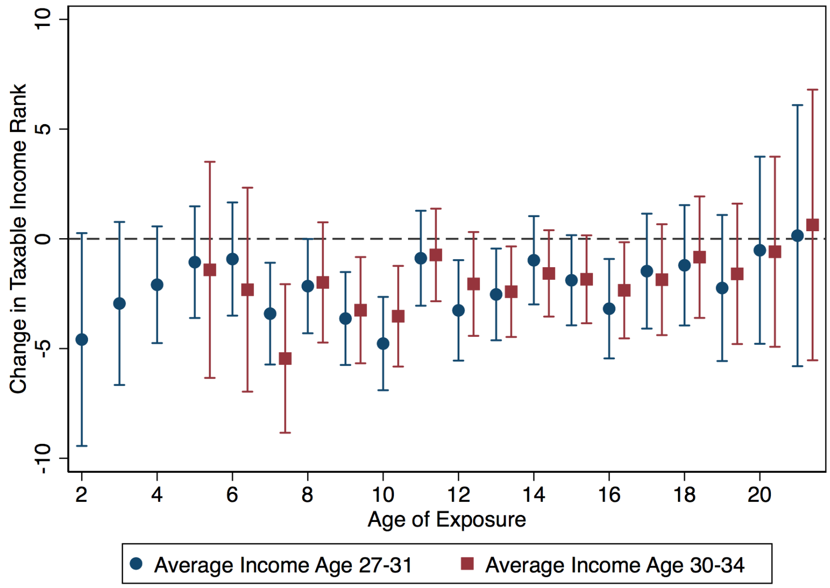


(a) Trees from the Random Forest

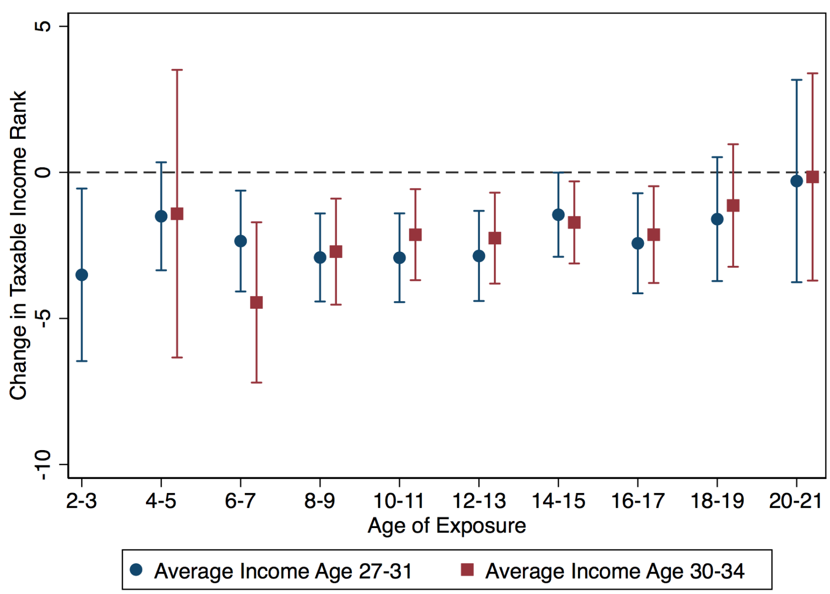
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	1	0	0	0	0	0.333	0	0
2	0	1	0	0	0.667	0	0	0
3	0	0	1	0.667	0	0.333	0.667	0.667
4	0	0	0.667	1	0	0.667	1	0.333
5	0	0.667	0	0	1	0	0	0
6	0.333	0	0.333	0.667	0	1	0.667	0.333
7	0	0	0.667	1	0	0.667	1	0.333
8	0	0	0.667	0.333	0	0.333	0.333	1

(b) Implied Proximity Matrix

Figure A2: Income Attainment Treatment Effects



(a) Reductions in Income Attainment by Age



(b) Reductions in Income Attainment by 2 Year Age Bins

This figure presents the average reduction in a child's taxable income rank from (i) age 27 to 31 and (ii) age 30 to 34 by her age of exposure to parental unemployment. Each point estimate corresponds to an estimate of of Δ_a from the regression presented in equation (4). Panel (a) presented estimates by age of exposure, while Panel B pools up to 2 ages of exposure together. In each regression, I control for the child's sex with a dummy variable which does not vary with the age of exposure (i.e. $\gamma_a = \gamma \ \forall a$). The error bars on each point estimate are 95% confidence intervals, with standard errors clustered at the match-pair level. All estimates from the IID are adjusted to satisfy Statistics Canada's vetting rules, ensuring the privacy of individuals in the data.