Centre for Economic Performance

Discussion Paper ISSN 2042-2695

No.1733
Revised October 2021
(Replaced December 2020 version)

The grandkids aren't alright: the intergenerational effects of prenatal pollution exposure

Jonathan Colmer
John Voorheis

LSE The London School of Economics and Political Science

30 CEP YEARS

UKRI Economic and Social Research Council

Abstract
Using newly linked survey and administrative data, providing more than 150 million parent–child links, we show that regulation-induced improvements in early life air quality have intergenerational effects on human capital accumulation in the United States – the second-generation is more likely to attend college. Supporting evidence indicates that intergenerational transmission arises from greater parental resources and investments, rather than heritable channels. Our findings suggest that within-generation estimates of marginal damages substantially underestimate the total welfare effects of improving environmental quality.

Key words: air pollution, environmental regulation, social mobility, human capital
JEL codes: H23; Q53; J00

This paper was produced as part of the Centre's Education and Skills Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
We thank Anna Aizer, Doug Almond, Max Auffhammer, Alan Barecca, Andrew Barr, Spencer Banzhaf, Raj Chetty, Janet Currie, Lucas Davis, Chloe East, James Fenske, Leora Friedberg, John Friedman, Chloe Gibbs, Michael Green-stone, Nathaniel Hendren, Bill Johnson, Nicolai Kuminoff, Jason Lindo, Lee Lockwood, Matthew Neidell, Marianne Page, Maya Rossin-Slater, Chris Ruhm, James Sallee, Nicholas Sanders, Jay Shimshack, Bill Shobe, Jeffrey Shrader, Sandip Sukhtankar, Chris Timmins, Sarah Turner, Reed Walker, Randall Walsh, and Catherine Wolfram for many helpful thoughts, comments, and discussions. We are also grateful to seminar partic-ipants at Arizona State, UC Berkeley, the University of Birmingham, Cambridge, Columbia, the University of Exeter, Georgetown University, Harvard University, the Heartland Environmental & Resource Workshop at Illinois, Imperial College London, the London School of Economics, the University of Oregon, the University of Oxford, the NBER Children's Spring Meeting, UC San Diego, Santa Clara University, the University of Virginia, the U.S. Census Bureau, and the University of Warwick for helpful comments and suggestions. All errors and omissions remain our own. This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The statistical summaries reported in this paper have been cleared by the Census Bureau's Disclosure Review Board, release authorization numbers CBDRB-FY18-232, CBDRB-FY18-248, CBDRB-FY18-287, CBDRB-FY19-099, CBDRB-FY19-136, CBDRB-FY19-261, CBDRB-FY2020-CES010- 001, CBDRB-FY2020-CES010-015, CBDRB-FY2020-CES010-017, CBDRB-FY2020-CES010-029, CBDRB-FY2020-CES010-031, CBDRB-FY2020-CES010-001, CBDRB-FY21-CES014-004, CBDRB-FY2021-CES010-005, CBDRB-FY2021-CES010-012, CBDRB-FY2021-CES010- 030 and CBDRB-FY2021-CES010-031.

Jonathan Colmer, University of Virginia and Centre for Economic Performance, London School of Economics. John Voorheis, U.S. Census Bureau.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

When the National Ambient Air Quality Standards were introduced in the United States as part of the 1970 Clean Air Act Amendments, concerns about air pollution were largely focused on respiratory health and visible air quality. In the decades since academics and policymakers alike have learned that the impacts of air pollution are more far-reaching, affecting many dimensions of health, education, economic productivity, and overall well-being (Graff Zivin and Neidell, 2013; Chay and Greenstone, 2005; Ebenstein et al., 2016; Isen et al., 2017; Voorheis, 2017; Deryugina et al., 2019).

Children are especially vulnerable to the effects of air pollution. Growing evidence suggests that exposure to pollution and other environmental risks in early childhood can play a critical role in shaping economic opportunity, with persistent effects on health and well-being (Chay and Greenstone, 2003a,b; Almond, 2006; Black et al., 2007; Currie et al., 2009; Kelly, 2011; Almond et al., 2010; Isen et al., 2017; Black et al., 2018; Gronqvist et al., 2020). Given that parental income and human capital are considered important for the economic opportunities of children (Becker and Tomes, 1979; Cunha and Heckman, 2007; Heckman and Mosso, 2014; Lee and Seshadri, 2019), alongside documented high levels of intergenerational persistence (Solon, 1992; Black et al., 2005; Black and Devereux, 2011; Chetty et al., 2014; Chetty and Hendren, 2018a,b; Chetty et al., 2018, 2019; Black et al., 2020), this raises an important question: do environmental shocks to endowments at birth have persistent effects across generations? The answer to this question is key to understanding the role that environmental quality might play in shaping broader patterns of economic opportunity and social mobility, as well as having important implications for evaluating the returns to investments in environmental quality.

We provide new evidence that prenatal particulate matter exposure has intergenerational consequences. To identify this effect we exploit the introduction of the 1970 Clean Air Act Amendments (CAAA), which imposed county-level limits on the maximum-allowable concentrations of total suspended particles (TSP), resulting in substantial improvements in ambient air quality in hundred of counties across the United States. This variation has previously been used to explore

the effects of air pollution exposure on contemporaneous outcomes such as infant mortality (Chay and Greenstone, 2003a), home prices (Chay and Greenstone, 2005), fetal mortality (Sanders and Stoecker, 2015), adult mortality (Chay et al., 2003), and later-life earnings for those that were born at the time of the 1970 CAAA (Isen et al., 2017). The existing literature focuses on people that were directly affected by the 1970 CAAA. However, if regulation-induced reductions in pollution exposure have intergenerational effects, then within-generation analyses may substantially underestimate the total welfare effects of investments in environmental quality.

Thanks to a new set of parent–child linkages constructed using the U.S. Census Bureau's data linkage infrastructure (combining Decennial Census data, household surveys and administrative records), we are able to explore the intergenerational consequences of improvements in environmental quality. Our data exploits the exact date and location of birth for all children born in the 1960s and 1970s and more than 150 million parent–child links.

We focus on the second generation — the children of those who directly benefited from the 1970 CAAA. We analyze the difference in outcomes for the children of cohorts born in counties before and after improvements in air quality to the children of those cohorts born in counties that didn't experience improvements. A consequence of this research design is that all cohorts should experience the same air pollution later in life. Children born before the Clean Air Act experience the reduction in pollution once the regulation comes into effect. This helps to isolate the additional effect of parental exposure to clean air in very early childhood relative to improvements in the air quality of parents born before the introduction of the 1970 CAAA. In support of this assumption we estimate that, on average, there are no differences in later-life particulate matter exposure for the first generation, nor any differences in prenatal or later-life particulate matter exposure for the second generation.

Our results suggest that prenatal exposure to particulate matter has a statistically significant and economically meaningful impact, not only on those directly affected, but also on the economic opportunities of their children. We estimate that children whose parents experienced lower gestational pollution exposure are more likely to attend college. A $10\mu g/m^3$ reduction in parental

gestational TSP exposure is associated with a 2.27–2.51 percentage point increase in the likelihood of attending college.[1] This is equivalent in magnitude to the intergenerational effects of 0.135-0.15 Head Start Programs (Barr and Gibbs, 2017).

In light of this effect, we investigate the mechanisms that may drive intergenerational transmission. We posit that there are two broad channels through which regulation-induced improvements in parental health could affect the second generation. The first channel is biological. Prenatal pollution exposure could result in epigenetic changes – that is, permanent changes in gene expression – that are hereditarily transmitted to the second generation. The second channel is economic. Prenatal pollution exposure affects the health, human capital, and earnings of the first generation. In turn, parental resources and investments may affect the likelihood of their children attending college.

First, we explore the empirical relevance of the biological pathway. We do so using information on whether parent-child links reflect adopted, biological or step-children. We estimate that the non-biological children of affected parents are also more likely to attend college. If anything, we find a larger effect for non-biological children, indicating that a heritable epigenetic mechanism is unlikely to be a first-order driver of intergenerational transmission. We caveat that other socio-economic differences between biological children and non-biological children, or socio-economic differences between the parents of adopted/step-children may also contribute to the differences in the estimated effects and so we cannot rule out the existence of a heritable transmission channel. There may also be latent health effects associated with inherited epigenetic changes that are not realized at this stage of the life cycle. Nevertheless, the results suggest that economic factors rather than heritable epigenetic mechanisms appear to be of first-order importance in explaining the college attendance effect.

Second, we explore whether intergenerational transmission might be driven by economic forces,

[1]The implied estimate from the reduced-form effect of the 1970 CAAA, which, on average, induced an $11.9 \mu g/m^3$ reduction for the parents of the second-generation sample, is almost identical. An 11.9 $\mu g/m^3$ reduction in parental gestational TSP exposure is associated with a 2.7-2.98 percentage point increase in the likelihood of attending college. Our reduced-form estimates range from 2.72-2.98 percentage points. In proportional terms, an 11.9 $\mu g/m^3$ decrease in TSP exposure from the 1971 baseline would be similar to the cross-sectional difference in the Black-White particulate matter exposure gap in the early 2000s (Currie et al., 2020).

that might affect parental resources and investments. We estimate that a $10\mu g/m^3$ reduction in prenatal TSP exposure resulted in a 0.7 percent increase in annual earnings. This translates to a $2,705 increase in cumulative lifetime earnings in present value terms (a 5% discount rate).[2] While this effect is non-trivial, the existing literature suggests that the effect size is likely too small to plausibly explain the entirety of our college-attendance estimate (Lovenheim, 2011; Lovenheim and Lockwood Reynolds, 2013; Bulman et al., 2017).[3] Taking the largest estimates from the existing literature our earnings effect can account for as much as 20% of the overall college attendance effect if both parents were exposed.

We also explore the degree to which differential selection into our second-generation sample could drive our effects. This would be an intergenerational transmission mechanism at the intersection of biology and economics. Based on existing evidence, the cohorts that experienced lower prenatal particulate matter exposure end up healthier and richer than cohorts that were born prior to the introduction of the 1970 CAAA. These factors may differentially affect fecundity or fertility choices (Becker, 1960; Lovenheim and Mumford, 2013; Black et al., 2013; Dettling and Kearney, 2014; Kearney and Wilson, 2018; Carre et al., 2017; Clay et al., 2018; Autor et al., 2019). If treated individuals have fewer children, then the college attendance effect could be driven by selection. We estimate precise null effects of prenatal pollution exposure on the likelihood of having children, the number of children, or the age at which parents were when they had their first child. We do not estimate any differential effects of prenatal exposure on fertility outcomes for men and women.

In terms of family structure, we estimate that lower exposure to prenatal pollution is associated with a decrease in the likelihood of ever being married, but also a decrease in the likelihood of getting divorced conditional on getting married. This suggests that the quality of marriages within treated cohorts is higher. However, the effects are small, corresponding to a 0.4/0.3 percentage

[2]This estimate is slightly smaller than the estimates provided by Isen et al. (2017), who examine the effects of the 1970 CAAA on first generation later life earnings at age 30 using the LEHD for 24 states.

[3]We note that the existing literature explores the effects of an increase in individual wealth much later in life (usually during the teenage years of the second generation). In our context, the second generation benefit from increased parental income for their entire life, including the early years where increased household resources may be particularly important (Bastian and Michelmore 2018, Bailey et al. 2020) The effect of household wealth in early childhood on college attendance would have to be approximately three times larger than the effects of household wealth on college attendance in adolescence for earnings to fully explain our result.

point decrease in the likelihood of ever being married/divorced for each 10 $\mu g/m^3$ reduction in prenatal particulate matter exposure. While previous research suggests that the children of divorce may be substantially less likely to attend college (Painter and Levine, 2000; Gruber, 2004; Frimmel et al., 2016), the magnitude of the effect suggests that divorce could only explain up to 1.5% of our result if we assume that the children of divorce are 10 percentage points less likely to attend college.

Another possible mechanism is that healthier parents lead to greater parental investments in cognitive and non-cognitive skills, which have long been considered important inputs to educational and labor market outcomes (Heckman and Rubenstein, 2001; Cunha and Heckman, 2007; Weinberger, 2014; Deming, 2017). In terms of their effect on college attendance, we calculate that a 10 $\mu g/m^3$ reduction parental prenatal exposure to particulates would be equivalent to endowing their child with 0.135-0.15 Head Start Programs (Barr and Gibbs, 2017), a 1.2 standard deviation increase in self esteem, a 0.42 standard deviation reduction in impulsivity, a 0.73 standard deviation decrease in schooling problems, or a 0.23 standard deviation increase in cognitive ability (Lundberg, 2017). While it is unlikely that any of these considerations are likely to drive our college attendance result by themselves, a combination of these considerations could be affected by parental investments as a residual explanation for our findings.

To provide a more concrete understanding of whether parental investments in cognitive and non-cognitive skills is a plausible mechanism, we link the American Time Use Survey (ATUS) to administrative records on place of birth. Following (Price, 2008) we calculate the amount of "quality time" that parents spend with their children. We estimate that a 10 $\mu g/m^3$ reduction in parental prenatal pollution exposure is associated with an additional 25 minutes of quality time spent with children each day, a 20% increase relative to the control mean. These findings suggest that parental investments in cognitive and non-cognitive skills may contribute to the estimated college attendance increases. We caveat that the sample size is substantially smaller in this analysis.

Overall, our results suggest that the intergenerational transmission of parental prenatal pollution exposure likely reflects a combination of parental resources and investments, which may have

been shaped in part by a more secure family environment. Those directly affected by lower prenatal pollution exposure are richer and healthier. Consistent with earlier study, the parental earnings effects do not appear to be driven by educational attainment and so may reflect broader improvements in health that may have had direct effects on the second generation through increased "parenting human capital."

Our findings contribute to several literatures. First, we contribute to the literature documenting the importance of environmental factors, rather than genetic factors, in determining human capital endowments at birth (Chay and Greenstone, 2003a,b; Almond, 2006; Black et al., 2007; Currie et al., 2009; Fertig and Watson, 2009; Kelly, 2011; Almond et al., 2010; Isen et al., 2017; Black et al., 2018; Gronqvist et al., 2020). Specifically, we contribute to a nascent literature documenting persistence in the economic consequences of early childhood and prenatal shocks across generational boundaries (Black et al., 2018; Barr and Gibbs, 2017; East et al., 2017; Nilsson, 2017; Akresh et al., 2018). Our estimates suggest large intergenerational benefits from reductions in air pollution. We build on this literature by exploring the mechanisms that could underlie the transmission of such shocks. Parental resources and investments appear to drive our result.

Second, we contribute to what is now a well-established literature on the Clean Air Act (Chay and Greenstone, 2003a; Chay et al., 2003; Sanders and Stoecker, 2015; Isen et al., 2017). Our results highlight the intergenerational benefits of environmental regulations, with important implications for human capital accumulation. Back-of-the-envelope calculations suggest that the combined total first-generation and (implied) second-generation discounted lifetime earnings effect is 50% larger than solely focusing on the first-generation earning effects. This combined total accounts for 55-70% of the monetized damages associated with infant mortality examined in previous work (Chay and Greenstone, 2003a).[4] To date, the mortality effects of pollution have been the largest monetized damage in formal benefit–cost evaluations of environmental regulations. Mortality benefits currently account for 93% of EPA calculated benefits. As such, the absence of long-run and intergenerational considerations in benefit–cost analysis suggests that

[4]The range of estimates arises because we bound the value of a statistical life between $9 million and $11.75 million ($2016), following Lee and Taylor (2019)

the returns to improvements in environmental quality may be substantially larger than previously thought. These effects are important but challenging to measure due to the data requirements. How can we account for these effects in contexts where administrative data linkages are not available? Since our findings suggest that intergenerational transmission is driven by economic factors, well-identified first-generation income effects combined with available estimates of intergenerational income elasticities could be used to provide an (appropriately discounted) lower-bound estimate of second-generation effects in other contexts.

Third, we contribute to the literature on the role of parental resources and investments in shaping child outcomes Becker (1981); Cunha and Heckman (2007); Bjorklund and Salvanes (2011); Doepke and Zilibotti (2017); Doepke et al. (2019). Our findings suggest an important role for parental-child interactions, as well the role that greater parental income in early childhood may play, in shaping college attendance. While descriptive evidence on the potential role of these factors has long been available, identifying exogenous variation has been challenging. By exploiting plausibly exogenous variation in parental endowments at birth we argue that our result provide compelling, though indirect, evidence on the importance of parental health, wealth and parent-child interactions on college attendance.

Finally, we contribute to the literature exploring economic mobility and opportunity (Becker and Tomes, 1979; Solon, 1992; Black et al., 2005, 2007; Black and Devereux, 2011; Chetty et al., 2014; Chetty and Hendren, 2018a,b; Chetty et al., 2018). Within this literature, the importance of neighborhoods has been established for the economic opportunities of children (Chetty et al., 2016; Chetty and Hendren, 2018a; Chyn, 2018; Deutscher, 2019). How neighborhoods matter is less clear. It is well established that economic and environmental inequality walk hand-in-hand. Individuals who live in low-income areas and disadvantaged communities are exposed to higher levels of pollution than those that live in high-income areas (Commission for Racial Justice, United Church of Christ, 1987; Mohai et al., 2009; Banzhaf et al., 2019; Colmer et al., 2020; Currie et al., 2020). Our findings suggest that poor environmental quality could have persistent effects on economic opportunity, raising new questions about the role that investments in environmental

quality might play in shaping upward mobility and economic opportunity, and inequality.

2 Background and Conceptual Framework

What are the channels through which prenatal pollution exposure could have intergenerational effects on human capital? First, we need to understand how prenatal particulate matter exposure could directly affect health and development. In our empirical analysis we exploit variation in prenatal exposure to total suspended particles (TSP), the type of particulate matter that was regulated by the EPA at the time of the 1970s Clean Air Act Amendments (CAAA). TSPs represent a complex mixture of organic and inorganic compounds found in the air that are smaller than 100 microns in diameter. TSPs enter the atmosphere from both human activity (e.g., traffic, construction, and industrial production) and natural sources (e.g., wildfire smoke, dust, pollen).

All particulate matter can damage human health, although larger particles are less harmful than smaller particles. This is because larger particles settle more quickly than smaller particles and so are less likely to be inhaled. When larger particles are inhaled, they settle in the nose and throat and are usually expelled by coughing or sneezing. By contrast, smaller particles (those smaller than 10 microns) can remain in the air for extended periods of time and, once they are inhaled, penetrate the lungs and the brain.

Once inside the body, particles are understood to affect respiratory function, lung development, and brain development. Since particles can pass from the lungs into the blood stream, they can also cause cardiovascular problems. These effects are amplified during gestation. Maternal exposure to particulate matter is understood to affect fetal health through a number of physiological pathways, such as oxidative stress, DNA damage to cells, damage to blood vessels, increased blood pressure, and reduced endothelial function (which could affect the transfer of nutrients to the fetus). Particles can also be transferred directly to the fetus through the bloodstream, affecting respiratory, neurological and cardiovascular development (Dejmek et al., 1999; Glinianaia et al., 2004; Bové et al., 2019; Huifeng et al., 2020).

Particulate matter is understood to affect human capital and development throughout the life cycle. Its effects are not limited to the gestational period, although marginal damages are understood to be larger during this period. This raises an important empirical challenge: we must isolate the effects of prenatal pollution exposure from contemporaneous effects experienced by the first generation throughout their life cycle, or by the second generation throughout their life cycle. Simply comparing the outcomes of individuals that experience prenatal exposure to locations with cleaner air to those of individuals who experience prenatal exposure to locations with dirtier air would not identify the effects of prenatal exposure because individuals born in "treatment" locations may be exposed to lower air pollution over their entire life cycle, and so may their children. To distinguish between the effects of prenatal exposure and cumulative exposure within and across generations, we require a research design that compares individuals who have different exposure to air pollution during gestation, but the same exposure post-birth. To formalize this idea we extend the framework presented by Isen et al. (2017).[5] In the process, we highlight the mechanisms through which prenatal pollution exposure could (1) affect the economic outcomes of those directly affected and (2) transmit across generations to affect their children.

Let an individual's health stock be a function of inputs during two time periods: $h = h(p_1, p_2, X(p_1))$, where p_1 represents prenatal pollution exposure, p_2 represents post-birth pollution exposure, and X reflects their genetic endowment, which may be affected by prenatal pollution exposure through epigenetic effects (permanent changes in gene expression). Our research design exploits variation in the introduction of the 1970s Clean Air Act Amendments (CAAA), which lowered TSP concentrations in counties that exceeded the air quality standards, following its implementation. We compare cohorts born just before and just after the CAAA came into effect in counties that were affected by the regulation to cohorts born just before and after in counties that were unaffected. In this analysis the treated cohort born after the CAAA is exposed to lower p_1 and p_2. By contrast, cohorts born in these counties before the CAAA came into effect had higher p_1 but the same exposure to p_2 (assuming they continued to live in their counties of birth). This comparison isolates the

[5]The foundations of this model are based on earlier work by Grossman (1972), Bleakley (2010), and Cunha and Heckman (2007).

additional effect of changes in prenatal pollution exposure p_1. More precise details on the research design and econometric specification are presented in section 3.

2.1 First-Generation Effects

For the first generation – those directly affected – we model the effects of prenatal pollution on earnings, which is a function of the health stock h and education e,

$$y_f = y_f(e_f, h_f) = y_f(e_f(h_f(p_{f,1}, p_{f,2}, X(p_{f,1}))), h_f(p_{f,1}, p_{f,2}, X(p_{f,1})))$$

where y_f represents earnings and e_f represents years of schooling. The channels through which prenatal pollution exposure affects earnings can be characterized by taking the total derivative of earnings, y_f, with respect to prenatal pollution exposure $p_{f,1}$,

$$\frac{dy_f}{dp_{f,1}} = \frac{\partial y_f}{\partial e_f}\frac{\partial e_f}{\partial h_f}\left[\frac{\partial h_f}{\partial p_{f,1}} + \frac{\partial h_f}{\partial X}\frac{\partial X}{\partial p_{f,1}}\right] + \frac{\partial y_f}{\partial h_f}\left[\frac{\partial h_f}{\partial p_{f,1}} + \frac{\partial h_f}{\partial X}\frac{\partial X}{\partial p_{f,1}}\right]$$

Prenatal pollution exposure, p_1, affects the health stock h, which affects earnings directly, $\frac{\partial y}{\partial h}$, as well as indirectly through educational attainment, $\frac{\partial y}{\partial e}\frac{\partial e}{\partial h}$.

2.2 Second-Generation Effects

For the second generation, we model the effects of first-generation (parental) prenatal pollution exposure on second-generation educational attainment, which is a function of the second generation's health stock h_s. The health stock of the second generation is a function of second-generation prenatal pollution exposure, $p_{s,1}$, second-generation later-life pollution exposure, $p_{s,2}$, first-generation earnings, $y_f(h_f, e_f)$, first-generation education, $e_f(h_f)$, first-generation health stock, $h_f(p_{f,1}, p_{f,2}, X(p_{f,1}))$, and any genetic changes that may have occurred due to first-generation prenatal pollution exposure or second-generation prenatal pollution exposure, $X(p_{f,1}, p_{s,1})$,

$$e_s(h_s(p_{s,1}, p_{s,2}, X(p_{f,1}, p_{s,1}); y_f(h_f(p_{f,1}, p_{f,2}, X(p_{f,1})), e_f(h_f(\cdot)); e_f(h_f(\cdot)); h_f(\cdot)))$$

The channels through which prenatal pollution exposure affects earnings can be characterized by taking the total derivative of educational attainment, e_s, with respect to parental prenatal pollution exposure $p_{f,1}$,

$$
\begin{aligned}
\frac{de_s}{dp_{f,1}} &= \frac{\partial e_s}{\partial h_s}\frac{\partial h_s}{\partial X}\frac{\partial X}{\partial p_{f,1}} && \text{(Biological Effects)} \\
&+ \frac{\partial e_s}{\partial y_f}\frac{\partial y_f}{\partial h_f}\left[\frac{\partial h_f}{\partial p_{f,1}} + \frac{\partial h_f}{\partial X}\frac{\partial X}{\partial p_{f,1}}\right] + \frac{\partial e_s}{\partial y_f}\frac{\partial y_f}{\partial e_f}\frac{\partial e_f}{\partial h_f}\left[\frac{\partial h_f}{\partial p_{f,1}} + \frac{\partial h_f}{\partial X}\frac{\partial X}{\partial p_{f,1}}\right] && \text{(Income Effects)} \\
&+ \frac{\partial e_s}{\partial e_f}\frac{\partial e_f}{\partial h_{f,1}}\left[\frac{\partial h_{f,1}}{\partial p_{f,1}} + \frac{\partial h_f}{\partial X}\frac{\partial X}{\partial p_{f,1}}\right] && \text{(Parental Education)} \\
&+ \frac{\partial e_s}{\partial h_f}\left[\frac{\partial h_f}{\partial p_{f,1}} + \frac{\partial h_f}{\partial X}\frac{\partial X}{\partial p_{f,1}}\right] && \text{(Parental Health)}
\end{aligned}
$$

Broadly speaking, parental prenatal pollution exposure could affect the educational attainment of the second generation through either biological channels, $\frac{\partial e_s}{\partial h_s}$, or the household environment, i.e, parental resources, $\frac{\partial e_s}{\partial y_f}$, and human capital, $\frac{\partial e_s}{\partial e_f}$ and $\frac{\partial e_s}{\partial h_f}$.

The remainder of this paper seeks to estimate $\frac{de_s}{dp_{f,1}}$, arising from changes in the level of parental prenatal particulate matter pollution exposure experienced at the time of the 1970 CAAA. We also analyze the mechanisms that help us to distinguish between biological channels and changes in the household environment. Note that an implicit assumption underlying the decomposition of these mechanisms is that $p_{f,2}$, $p_{s,1}$, and $p_{s,2}$ are not a function of $p_{f,1}$. If, for example, parental prenatal pollution exposure is positively correlated with pollution exposure in later life then we can't identify the effects of parental prenatal pollution exposure separately from later-life first-generation exposure or second-generation exposure. The following section provides more detail about the research design and empirical specifications that help us to isolate the effects of parental prenatal pollution exposure, as well as direct evidence that there is no differential effect of prenatal pollution exposure on later life pollution exposure; i.e, $\frac{\partial p_{f,2}}{\partial p_{f,1}} = 0$, $\frac{\partial p_{s,1}}{\partial p_{f,1}} = 0$, and $\frac{\partial p_{s,2}}{\partial p_{f,1}} = 0$.

3 Research Design

In this section we provide an overview of the data and sample construction process, as well as the empirical specifications used to estimate the effects of parental prenatal particulate matter exposure on second-generation outcomes.

3.1 Data

To study the intergenerational effects of pollution exposure, it is necessary to identify the location and date of birth for each parent, infer their exposure to ambient air pollution, link these parents to their children, and measure outcomes. No single dataset has all of these features, and so our analysis requires linking decennial Census, administrative records, and survey data. This linkage is done using unique anonymous personal identifiers called Protected Identification Keys (PIKs), which can be thought of as "scrambled" or hashed Social Security Numbers. PIKs are assigned to datasets using a probabilistic matching algorithm which links personally identifiable information (name, date of birth, Social Security Number, etc.) to a reference file of people in the United States.[6]

3.1.1 Parent–Child Links

We begin by assembling a database of parent–child links that can be evaluated using survey, decennial Census and administrative data sources available in the Census Bureau's data linkage infrastructure. We identify links in two main datasets: the full count decennial Census from 2000 and 2010, and the American Community Survey (ACS) from 2005–2015.[7] The set of links we are able to identify is not, we should stress, the full population of links. We will miss two main sets of parent–child linkages: parent–child linkages in households which formed and dissolved between decennial Censuses (who were not ACS respondents), and parent–child links in which either the

[6]For more on the process of PIK assignment see Wagner and Layne (2014) and Appendix C.

[7]Other probable links could in principle be measured in the universe of IRS 1040 tax returns. We focus on the Census and ACS links because we can infer with certainty that these links are between parents and their children.

parent or child cannot be assigned a PIK. We focus on two sets of these links: "certain" links, where the survey and Census information allows us to identify a parent–child link with certainty; and "probable" links, where the survey and Census information allows us to identify the spouse of a certain link parent, who are likely also parents (but may be step-parents).

Each dataset contains slightly different information on relationships within households. The decennial Census and ACS data both contain detailed information on relationships within the household, with one important limitation — the Decennial Census/ACS relationship question asks for information only on the relationship between an individual and the head of household. This means that we can identify parent–child links for the head of household with certainty. We additionally identify probable parent–child links between the head of household's married or unmarried partner and the head of household's children. For head of household–child links, we have additional information about the type of link — specifically whether a child is biological, adopted or a step-child. In the main analysis we focus on certain parent–child links to minimize any attenuation bias introduced through the incorrect assignment of children to affected, or unaffected, parents. In the appendix materials we show that our results are robust to incorporating probabilistic links.

3.1.2 Pollution Exposure

To analyze the intergenerational effects of pollution exposure, we need to infer the level of ambient air pollution that the first generation were exposed to during gestation. We do this in three steps. First, we link the set of unique parents identified above to the Census Numident to obtain date and place of birth. We then obtain monitor-level daily pollution measures from the EPA, which we aggregate to the county level, and link these county-level measures to the parents' place of birth.

The Census Numident is a person-level administrative records file derived from the Social Security's Numident, which contains all individuals who ever apply for a Social Security Number. The Numident contains information on individuals' exact date and place of birth. Since place of birth data are not standardized, we assign county of birth information to first and second generation individuals using a crosswalk provided by Isen et al. (2017), combined with a probabilistic match-

ing approach used in Voorheis (2017). For cases where county of birth cannot be assigned using the Isen et al. (2017) crosswalk, we match the Numident place of birth string to the United States Geological Survey's list of places using a fuzzy string matching algorithm based on the optimal string alignment (OSA) distance metric.

We infer the level of pollution exposure experienced by individuals based on average pollution concentrations during the 270 days of gestation within their county of birth. To gather this pollution exposure information, we rely on monitor data from the EPA, which we retrieve using a public facing API.[8] Our pollutant of interest is particulate matter. For the relevant period of time (around 1970) the primary regulatory definition of particulates was total suspended particles (TSP), defined as particulate matter with a density of less than 100 microns, measured in units $\mu g/m^3$. We retrieve data from all TSP monitors between 1969 and 1980.

The TSP standard was set based on a 24-hour sampling. Our baseline approach for aggregating the daily monitor-level observations is as follows: for each county-day, we calculate the average TSP concentration across all active monitors in that day, which we take as the average exposure to TSP in that county on that day. We then calculate county-level gestational period averages of prenatal TSP exposure for each birth between 1969 and 1980.

3.1.3 Outcomes

Finally, we require information on outcomes for parents and children, as well as other information on observable socio-demographic characteristics. We measure these outcomes using the ACS, which contains detailed information on family structure, human capital and labor market outcomes. Note that since the ACS is a nationally representative survey of a sample of households (about 1 percent per year), we observe outcomes from only a fraction of the parents and children identified above.

Our main outcome of interest is college attendance for the second generation. We measure college attendance contemporaneously – whether an individual was currently attending post-

[8] See https://aqs.epa.gov/aqsweb/documents/data_mart_welcome.html for more details.

secondary education at survey response. We focus on this outcome because we have limited data on earnings. Very few individuals over the age of 22 at ACS response have parents born after 1971, limiting the sample available to examine the effects on wages and labor force participation. In the absence of earnings data we focus on human capital accumulation, focusing on college attendance and high school completion. The reason we focus on human capital accumulation is because human capital measures available in the ACS are prevalent among people under the age of 22, who are much more likely to have parents that were affected by the 1970 CAAA. Additionally, although some second generation individuals would have earnings in IRS administrative records, there are very few, if any, individuals around the "overtaking age" when earnings are near the life-cycle peak.[9] Since we are measuring college attendance contemporaneously, our main sample for the college attendance analysis is ACS respondents aged 19-22 who can be linked to a parent born around CAAA 1970 for which we are able to measure pollution exposure.

We explore several outcomes for the first generation with a view to understanding the direct effects of prenatal pollution exposure and the mechanisms through which these effects could be transmitted across generations. Our main sample for the first generation analysis is all ACS respondents who were born between 1969 and 1980, for which we are able to measure pollution exposure and identify a parent-child link. Economic outcomes are explored by defining variables for unemployment, public assistance receipts, and wages from detailed ACS questions on income and labor force participation. We also explore family structure using information. In evaluating family size, we use the full sample of ACS respondents during our sample period, not just those for which a parent-child link was formed. We do this to explore selection into the second-generation sample — whether pollution exposure affects the likelihood of having children, the number of children, and conditional on having children the age at which the first child was born. We also look at family structure, i.e., marriage, divorce, and assortative matching measures, as we believe that these considerations could plausibly affect parental resources and investments. Finally, we look at how parents spend time with their children by linking the American Time-Use Survey with our existing

[9]For instance, in 2018, the only 30 year olds with parents born after the CAAA would have been born when their parents were 16.

data infrastructure. We caveat that sample size is substantially smaller in this analysis because we are restricted to ATUS respondents born in monitor-counties during our evaluation window. Following Price (2008) we construct a measure of "quality time", which restricts time spent with children to only include activities in which the child was the primary focus of the activity, or where a reasonable amount of interaction between parent and child would likely arise. This measure of quality time consists of 13 activities, which are split into 4 categories: Category A - reading to/with; playing, not sports; helping with homework; helping/teaching; talking with/listening to; arts and crafts; Category B - eating meals together; Category C - playing sports; attending performing arts; participating in religious activities; Category D - looking after (as primary activity); physical care.

3.1.4 Controls

In addition to our outcome variables, the ACS provides sociodemographic information, including details on race, sex, and age. The first and second generation do not always appear in the ACS at the same time, so we also collect demographic characteristics from the decennial Census to control for parental characteristics when evaluating second-generation outcomes. We also collect information on the characteristics of first- and second-generation county of birth — population, employment, personal income per capita and total transfer income — from the Bureau of Economic Analysis' Regional Product Accounts as control variables.

3.2 Empirical Strategy

The objective of our main empirical exercise is to estimate the relationship between an individual's prenatal particulate matter exposure and the later life outcomes of their children 40-50 years later. There are two key threats to identification. First, exposure to particulate matter is correlated with many observable and unobservable characteristics that are correlated with long-run economic and social outcomes, resulting in omitted variable bias. We also face a second identification challenge, which is specific to identifying the effects of early-life shocks on later life outcomes. We want to

isolate the effects of prenatal exposure. Even if the shock was as good as randomly assigned, one would want to show that prenatal exposure to a shock does not affect later life exposure. This was the focus of our discussion in the conceptual framework. If differences in early-life pollution exposure (or any early-life shock) are also correlated with differences in later-life pollution exposure, it is not possible to disentangle the persistent effect of early-life shocks from later-life exposures.

To address the endogeneity concerns related to pollution exposure and isolate the effects of early-life pollution exposure, we instrument for changes in particulate matter exposure using the introduction of the 1970 Clean Air Act Amendments (Chay and Greenstone, 2003a; Isen et al., 2017). The Clean Air Act was introduced in 1963 and regulates air pollution in the United States and is the largest environmental program in the country. It requires the EPA to develop and enforce regulations to protect the population from exposure to airborne pollutants that are known to be hazardous to human health. In 1970, the Clean Air Act was amended, authorizing federal regulations to limit emissions, resulting in a major shift in the federal government's role in air pollution control. As a consequence of the 1970 amendments the EPA established the national ambient air quality standards (NAAQS), specifying the maximum acceptable levels for six criteria air pollutants — sulfur dioxide (SO_2), particulate matter (TSP, $PM_{2.5}$ and PM_{10}), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3), and lead.

The NAAQS standards were applied at the level of Air Quality Control Regions (AQCRs), which were EPA-defined collections of counties. AQCRs that exceeded these pollution levels were considered to be nonattainment areas. As a consequence of being designated a nonattainment area, a plan must be implemented to meet the standard. The EPA defined an AQCR as being in nonattainment of NAAQS if any of the individual monitors in the AQCR fell outside the NAAQS standards. Counties with monitors out of attainment needed to make large improvements in air quality, while counties which were in attainment did not, even if they were located in an AQCR which was out of attainment (in other words, attainment counties in nonattainment AQCRs were inframarginal). The NAAQS standards are thus binding at the county level. The relevant fact for us is whether a county was in nonattainment of the NAAQS standards; however, the EPA did not pub-

lish county-level nonattainment designations until the late 1970s.[10] Thus we follow the existing literature in modelling the effect of the Clean Air Act using the nonattainment status of counties, not AQCRs, noting that it is the bindingness of EPA regulations, and not the nonattainment designations themselves, which produce our quasi-experimental variation in pollution exposure.

Following the existing literature we reconstruct nonattainment designations using TSP concentrations from the years before the CAAA was enacted. Nonattainment of the primary air quality standard for TSP set in the 1970 CAAA occur if either a county's annual average (geometric mean) TSP concentration is above 75 $\mu g/m^3$, or the second highest daily TSP concentration is above 260 $\mu g/m^3$. We use monitor-level observations to calculate the geometric mean and second highest daily TSP concentration for all counties with at least one monitor in 1970. This allows us to categorize 258 counties as "nonattainment" counties, and 319 counties as "attainment" counties.[11]

Following Chay and Greenstone (2003a) and Isen et al. (2017), we model the change in air pollution using an indicator variable for county nonattainment status interacted with an indicator for the years 1972 or later. The first-stage regression in this two-stage least squares estimator is a difference-in-differences regression model,

$$TSP_{j,c,t} = \alpha_0 + \alpha_1(Nonattainment_{c,1970} \times 1[\tau > 1971]) + \alpha_c^j + \alpha_{st}^j + \gamma X_j' + \delta X_c' t + \nu_{j,c,t} \quad (1)$$

Exposure is measured for parent j, where $TSP_{j,c,t}$ is the average particulate matter concentration that individual j was exposed to in county c and year t, measured in $\mu g/m^3$. $TSP_{j,c,t}$ is regressed on a time-invariant county indicator equal to 1 if a county is designated as nonattainment, $Nonattainment_{c,1970}$, and interacted with an indicator equal to 1 for the years after the CAAA went into affect, $1[\tau > 1971]$. The interaction term is therefore equal to 1 for nonattain-

[10] Note that, though in principle it would be possible to define an AQCR-level treatment, almost all monitored AQCRs had at least one monitor out of attainment, rendering this approach futile in practice.

[11] Consequently, we must restrict our analysis to first generation individuals born in these 577 counties and second generation individuals born to individuals born in these counties, since the pollution exposure of individuals born in other counties was unmeasured during this time period. Note, however, that these counties contained about two thirds of the US population in 1969 and cover all 50 states.

ment counties following the implementation of the 1970 CAAA. The parameter of interest is α_1, which provides a difference-in-differences estimate of the impact of nonattainment designation on prenatal TSP exposure in the years after CAAA regulations went into effect. Following Chay and Greenstone (2003a) and Isen et al. (2017) we also include additional controls. X_j' is a vector of individual characteristics, including age, race, and sex, as well as in-utero weather exposure. $X_c't$ is a vector of county-level characteristics, measured in 1969, interacted with linear and quadratic time trends. α_c^j are county-of-birth fixed effects that control for time-invariant unobserved determinants of the labor market outcomes and family structure for parents born in county c. α_{st}^j are birth-state $\times$ birth-year fixed effects which control for time-varying determinants of the long-run outcomes, common across all parents born in a state s in year t.[12] Across all specifications we cluster our standard errors by the first generation's county of birth — the level at which we measure exposure.

We show, consistent with previous research on the Clean Air Act, that the first stage relationship is strong — that nonattainment status is associated with significant and persistent declines in particulate matter concentrations in the years after the 1970 CAAA came into effect. Table 1, column 1, presents the first-stage relationship for the second-generation college attendance sample. We estimate that, on average, nonattainment is associated with an 11.90 $\mu g/m^3$ reduction in parental prenatal TSP exposure. In panel a) of Figure 1 we present cohort-specific estimates of the effect of nonattainment on prenatal TSP exposure using a distributed-lag model. Prior to the introduction of the 1970s CAAA we find no statistically significant, or economically meaningful, differences between nonattainment and attainment counties, providing support for the parallel trends assumption. Following implementation, we estimate an immediate and persistent reduction in prenatal TSP exposure. This is consistent with well-established findings in the existing literature (Chay and Greenstone, 2003a; Isen et al., 2017)

In addition to providing evidence in support of the parallel trends assumption, we show that there are limited differences between attainment and nonattainment county characteristics prior to the 1970 CAAA (Table A1). Isen et al. (2017) also show that a nonattainment designation is not

[12] In extensions we include second-generation controls, including second-generation birth-county fixed effects, α_c^i, second-generation birth-state $\times$ birth-year fixed effects, α_{st}^i, and second-generation individual characteristics, X_i'.

correlated with any changes to the observable characteristics of mothers that gave birth in the years following the 1970 CAAA.

We also present evidence in support of the assumption that the research design isolates the effects of prenatal pollution exposure. The introduction of the 1970s CAAA allows us to compare cohorts who have different TSP exposure during gestation, but the same exposure post-birth. Columns 2-4 of Table 1 show that there is no meaningful effect of parental prenatal exposure to nonattainment on first-generation later–life exposure (column 2), second-generation prenatal exposure (column 3), or second-generation later-life exposure (column 4). These findings are also supported in panels b), c) and d) of Figure 1. As with our first-stage estimate we estimate no meaningful differences prior to the introduction of the 1970 CAAA. Unlike the effects on prenatal exposure, we estimate no meaningful differences in later-life TSP exposure between nonattainment and attainment counties. This supports the premise of the research design, that the children born in nonattainment counties before the 1970 CAAA came into effect still benefited from cleaner air later in life. As such, we are confident that the research design isolates the differential effect of early life exposure to TSP.[13]

In the second stage, we use the predicted TSP levels from equation 1 in place of observed TSP levels,

$$Outcome_{i,j,c,t} = \rho_0 + \rho_1 \widehat{TSP}_{j,c,t} + \alpha_c^j + \alpha_{st}^j + \gamma X_i' + \delta X_c' t + \varepsilon_{i,j,c,t,y} \tag{2}$$

Our specification is the same as Chay and Greenstone (2003a) and Isen et al. (2017), except that our outcomes are measured for child i born to parent j. The coefficient of interest, ρ_1, captures the effect of a one-unit increase in CAAA-driven first-generation prenatal TSP exposure on the second-generation child i's later life outcomes.

The main concern regarding the instrumental variable research design is the exclusion restriction. To identify the intergenerational effects of TSP exposure it must be the case that the CAAA

[13] In Table A2 we also show that there is little evidence of differential migration, a plausible mechanism through which later-life differences in pollution could arise.

does not affect outcomes in any other way than through reductions in pollution. Isen et al. (2017) make the point that nonattainment designations could affect economic competitiveness (Greenstone, 2002; Greenstone et al., 2012; Walker, 2011, 2013). However, existing evidence suggests that the effects on the broader local economy are small, affecting less than 0.7 percent of the total workforce and is based on the 1990 Clean Air Act Amendments (Walker, 2013). Since the 1970 CAAA was the first major regulation to be introduced, actions to reduce emissions may have been less costly in 1970 than in the 1990s, there being more low-hanging fruit, attenuating concerns about economic competitiveness. Nevertheless, it is possible that the 1970 CAAA contributed to a decline in economic conditions for nonattainment counties, affecting the long-run economic prospects of affected individuals. Since effects on competitiveness would be expected to have the opposite effects on health to reductions in pollution exposure, our 2SLS estimates will understate the intergenerational effects of pollution exposure if the exclusion restriction is violated. The reduced form effect of nonattainment remains valid and is interpreted as the intergenerational effects of the 1970 Clean Air Act Amendments, rather than pollution. Our reduced form and corresponding 2SLS estimates produce conceptually identical results.

4 Results

Table 2 presents the results of our analysis on the effects of first-generation prenatal pollution exposure on the likelihood that the second-generation attends college, using our preferred second-generation sample that uses certain parent–child links. We estimate that a 10 $\mu g/m^3$ reduction in first generation prenatal pollution exposure is associated with a 2.27–2.51 percentage point increase in the likelihood that the second generation attends college.[14] Relative to the mean, this corresponds to an 5.7% increase in college attendance. This is a substantial effect, equivalent in magnitude to the second-generation effect of 0.135–0.15 Head Start Programs (Barr and Gibbs,

[14]In Table A3 we estimate the effects for maternal exposure, paternal exposure separately. We estimate slightly larger effects for paternal exposure, and the largest effect if both parents were exposed, however, the combined effect is smaller than the sum of the mother-only and father-only effects.

2017). In Panel B of Table 2 we present reduced form effects, capturing the effect of the 1970 CAAA on second-generation outcomes. We estimate that the reduced form effect of the 1970 CAAA is associated with a 2.72–2.98 percentage point increase in the likelihood of college attendance. Using the average nonattainment induced changes in TSP for this sample ($11.9\mu g/m^3$) our IV estimates imply a 2.7–2.98 percentage point increase in the likelihood of attending college. The 2SLS and reduced form estimates are almost identical. Figure 2 presents cohort-specific estimates of the reduced form effect of nonattainment on second-generation college attendance. Similar to the pollution outcomes, we do not estimate any differential pre-treatment effects. Following the introduction of the 1970 CAAA we estimate an immediate and sustained 2–3 percentage point increase in college attendance for the 1972 cohort onwards, however, similar to Isen et al. (2017) the cohort-specific estimates are more noisily estimated than the pooled estimates.

While we are not able to examine the intergenerational effects of prenatal pollution exposure on earnings at this stage, the effects on college attendance plausibly have meaningful impacts on the earnings potential of second-generation individuals. If the college wage premium is $28,000 p.a. (estimated using the public use CPS-ASEC), then combined with a graduation rate of 50 percent to convert college attendance to college completion[15], and a 3 percent real discount rate (5 percent discount rate + 2 percent wage growth) we predict a $987/$10\mu g/m^3$ increase in second generation cumulative lifetime earnings, discounted to age zero of the first generation. Combining the number of children born in nonattainment counties in 1972 (1.52 million), the probability that one of these individuals had a child (0.63), and the average number of children that were born to these individuals (2.24), we calculate that the total second generation earnings effects for the 1972 cohort is around $2.1 billion dollars. If reductions in TSP were persistent (and evidence suggests they were) we would expect these benefits to be realized for each cohort since 1972. If the intergenerational effects of prenatal pollution exposure have effects on income other than through college, this number represents a lower-bound estimate of the second-generation benefits.

We note an important caveat that our estimates are based on a selected sample. The children in

[15] See https://nces.ed.gov/programs/coe/indicator_ctr.asp

our second-generation analysis sample are aged 19-22 and so are more likely to be the first-born children of relatively younger parents. This may affect the degree to which our results generalize to all second-generation individuals. However, it is not possible to sign the selection bias. First-born children may be more likely to attend college dampening the marginal effect for this sample. Children born to younger parents may be less likely to attend college increasing the marginal effect for this sample. We note that reducing the time-window over which we estimate the effects (reducing the share of first-born children born to younger parents in the sample) has very little effect on our estimates (Tables A4).

Our results are also qualitatively robust to including probabilistic links, although the estimates are smaller in magnitude. This is consistent with the introduction of measurement error that would arise through the introduction of false links (Table A5).

We also conduct a permutation inference exercise, to address any concerns that the results are driven by sampling variability or spurious trends. Such concerns may be reasonable in the context of exploring long-run or intergenerational effects, where it is possible that even small amounts of confounding residual variation could be propagated over time if serially correlated, resulting in a spurious divergence of outcomes between treatment and control groups. Holding the sample fixed we randomly re-assign nonattainment status across counties 10,000 times and use these placebo realizations of nonattainment to estimate the original model. We then compute the share of placebo estimates that are higher in absolute value than our original estimate and calculate an exact p-value. For the association between first-generation prenatal nonattainment exposure and first-generation prenatal TSP exposure (our first-stage) we calculate an exact p-value of 0.000. For the association between first-generation prenatal nonattainment exposure and second-generation college attendance we calculate an exact p-value of 0.0136. These findings provide provide compelling evidence that our estimates are not driven by sampling variability or spurious trends.

Finally, we look at the likelihood that this sample drops out of high school as an additional measure of educational attainment (A6). We do not estimate any statistically significant effect on

the likelihood of high school completion, but note that dropping out of high school is a relatively rarer event than attending college and so may be less responsive to any channels driving college attendance.

4.1 Exploring Mechanisms

What are the mechanisms through which prenatal pollution exposure propagates across generations? We posit that there are two broad channels through which these effects could arise. The first channel is biological. Gestational pollution exposure could result in epigenetic changes, i.e., permanent changes in gene expression. If so, hereditary transmission of these changes could affect the second generation. The second mechanism is economic. Gestational pollution exposure affects the human capital and earnings of the first generation. In turn, parental resources and investments may affect their children's human capital accumulation.

4.1.1 Biological Effects

To explore the empirical relevance of the hereditary mechanism we examine the differential effect of first-generation prenatal pollution exposure between biological and non-biological (adopted and step) second-generation children. This exercise allows us to examine the degree to which permanent changes in gene expression, triggered by prenatal pollution exposure, are passed down from parent to child, affecting human capital, and in turn the likelihood of, and returns to, attending college. As the Decennial Census and ACS both ask whether the child of the head of household is biological, adopted, or step-child, we are able to identify a set of parent–child links for which there should be no direct hereditary mechanism.

Biological children could be affected through both changes in household environment and through genetic channels. By contrast, adopted and step-children should only be affected through the household environment (i.e., parental resources and investments).[16] If the effects on college attendance are entirely driven by genetic pathways then we would expect there to be no effect on

[16] We caveat that we do not know whether the biological parents of adopted children were treated.

adopted children. The differential effect would be the inverse of the effect on biological children. If the college attendance effect is entirely driven by the household environment then there should be no differential effect on adopted children, and the coefficient on the interaction term should be zero.

Of course, it is entirely possible that the college attendance effect is a mix of both channels, at which point the effect on adopted children should be smaller than the effect on biological children. Finally, we acknowledge that if differences between biological and non-biological children or biological and non-biological parents result in differential investments in adopted or step-children, the effect could be larger than the effect on biological children. In such a case, we can't rule out the presence of hereditary effects, but note that in such a case non-hereditary channels would appear to be relatively more important.

There are three potential concerns with this exercise. First, there may be a differential propensity to have non-biological children in nonattainment counties, resulting in selection into the second-generation sample. For example, if prenatal exposure to pollution reduces fertility or fecundity, the children of parents who benefited from reductions in prenatal pollution may be more likely to be biological. In column 1 of Table 3 we estimate that there is no statistically significant or quantitatively meaningful effect of prenatal TSP exposure on the likelihood that the child is adopted or a step-child, suggesting that this is not a first-order concern. Our second concern is that adopted children may be affected through hereditary channels if their biological parents were born in nonattainment counties at the same time as their adoptive parents. However, this requires that their adoptive parents and biological parents are born at the same time and location, which we argue is unlikely to be the case in a systematic way. In support of this assumption, we observe that adopted children are more likely to be born in a different county from their adoptive parents than biological children. Finally, implicit in this analysis is that non-biological and biological children/parents are the same in other dimensions. To the degree that biological children are more likely to attend college our estimates will be biased in favor of a hereditary channel. To the degree that biological children are less likely to attend college our estimates will be biased in favor of eco-

nomic channels. We note that there is no difference in the likelihood of attending college between biological and non-biological children in the control group.

Columns 2 and 3 of Table 3 estimate the effect of parental prenatal TSP exposure on second-generation college attendance for biological and non-biological children. We estimate that lower prenatal pollution exposure is associated with significant increases in the likelihood of attending college for both biological and non-biological children. This suggests that the estimated college attendance effect likely arises through parental resources and investment channels, rather than a hereditary mechanism (Bjorklund and Chadwick, 2003; Bjorklund, 2006; Bjorklund et al., 2007, 2010; Black et al., 2020; Kleven et al., 2020). We note that the estimated effects for non-biological children are larger than the effects for biological children. As such, we caveat that heterogeneity in the college attendance effect driven by other differences between biological and non-biological children/parents may confound this interpretation. Given that such differences likely reflect socio-economic factors, we argue that this provides further support for the premise that non-hereditary mechanisms are of first-order importance.

4.1.2 Income Effects

Turning to economic channels, we explore the effects of prenatal pollution exposure on first-generation labor market outcomes. Table 4 presents the results of this analysis.

We estimate that a 10 $\mu g/m^3$ reduction in in-utero exposure to TSP is associated with a 0.7 percent increase in earnings at the time of ACS response. Evaluated at the mean earnings, this corresponds to an annual earnings effect of $284 compared to the control mean. This effect is slightly smaller than the estimates presented in Isen et al. (2017). Differences in these effects are likely driven by differences in the age of evaluation, as well as the quality of the earnings data. Isen et al. (2017) estimate effects at age 29-31 using data from the 24 LEHD states. Our sample contains parents in all states which had active particulate monitors in 1969, and is looks at self-reported earnings later in life (after peak earnings).

Assuming that the estimated earnings effect is constant over the life cycle, and that earnings

are discounted at a real rate of 3 percent back to age zero, we calculate that the lifetime earnings effect of a 10 $\mu g/m^3$ increase in TSP is \$2,705 per person (\$4.1 billion for the 1972 cohort).

We also explore effects on labor force participation (column 2), unemployment (column 3), and public assistance (column 4). We fail to reject the null hypothesis that there is no relationship between gestational TSP exposure and these outcomes. The absence of any effects suggest that TSP exposure does not appear to affect labor market outcomes on the extensive margin. This is consistent with Isen et al. (2017) who estimate increases in the number of quarters worked, rather than extensive margin labor force participation effects. In Appendix A.4 we explore first-generation educational outcomes. We fail to reject the null hypothesis that prenatal TSP exposure has any effect on years of schooling or college attendance (Table B1). This is consistent with Isen et al. (2017) who argue that the effects on earnings are too small to be driven by schooling. Collectively, these results suggest that the earnings effects likely reflect a broader health sequelae.

Figure 3 presents cohort-specific estimates that are consistent with the pooled estimates in Table 4. We do not estimate any meaningful differential trends prior to the introduction of the 1970 CAAA. This provides further support for the parallel trends assumption.

Existing estimates of the relationship between parental wealth and college attendance suggest that our college attendance effect is unlikely to be driven solely by parental resources. Lovenheim (2011) and Lovenheim and Lockwood Reynolds (2013) explore the effects of an increase in household wealth on college attendance finding that a \$10,000 increase in household wealth is associated with a 0.71–0.92 percentage point increase in the likelihood of attending college. Bulman et al. (2017) explore the effects of winning the lottery on college attendance. They find that a \$10,000 increase in household wealth is associated with a 0.2 percentage point increase in attending college. If we take the largest estimate from the existing literature, assume that both parents were affected, and that our findings were driven entirely by the increase in one parent's earnings then we would expect a 0.48 percentage point increase in the likelihood of attending college. This would account for 20% of our estimate.

There are differences in our context, compared to the existing literature, however, that may

provide support for income as an important driver. In the existing literature potential college attendees are usually in their teenage years at the time that the wealth shock is realized. For example, Lovenheim (2011) and Lovenheim and Lockwood Reynolds (2013) look at wealth shocks in the 4 years prior to turning 18, and most of the first born children in Bulman et al. (2017) are teenagers at the time that their parents win the lottery. It is entirely possible that wealth shocks during the teenage years have much less of an effect on college attendance since many important human capital decisions have already been realized. Drawing on insights from the early childhood literature it is plausible that parental wealth shock in early childhood could have a larger effect on college attendance. Nevertheless, the effect of parental wealth in early childhood on college attendance would have to be more than five times larger than the existing literature to fully reconcile our findings. While parental earnings likely play some role it does not appear likely that the college attendance effect is fully explained by the earnings effect alone.

4.1.3 Family Structure and Fertility Effects

Another possibility is that part of the college attendance effect reflects selection into the second-generation sample through fertility effects. If those that were exposed to lower levels of pollution are richer and healthier this may affect the likelihood of family formation, through marriage and/or fertility choices. Becker (1960) posits that wealthier parents may have a preference for quality over quantity of children. However, empirical evidence suggests that increases in earnings or wealth are associated with increases in fertility (Lovenheim and Mumford, 2013; Black et al., 2013; Dettling and Kearney, 2014; Kearney and Wilson, 2018; Autor et al., 2019). If, despite the existing literature, parents that were exposed to lower levels of TSP have smaller families then under the plausible assumption that family size is negatively correlated with the likelihood of college attendance then this could explain part of our results.

Health and income differences may affect the likelihood of family formation, affecting the likelihood that the first generation have children, or the household environment in which children are born. It is also possible that pollution exposure could affect fertility or fecundity (Carre et al.,

2017; Clay et al., 2018). In such a case there may be reduced selection into the second-generation sample for the control group.

Table 5 presents our findings, exploring the fertility and family structure effects of prenatal exposure to TSP. We fail to reject the null hypothesis that there is no difference in the likelihood of having children, the number of children, or the age of parents when they had their first child. We estimate small negative effects on the likelihood of getting married, but also on the likelihood of getting divorced, conditional on marrying. This may point to an increase in the quality of marriages, resulting in a more secure family environment, however, the magnitude of the effects are small, corresponding to a 0.4/0.3 percentage point decrease in the likelihood of ever being married/divorced for each 10 $\mu g/m^3$ reduction in prenatal particulate matter exposure. While previous research suggests that the children of divorce may be substantially less likely to attend college (Painter and Levine, 2000; Gruber, 2004; Frimmel et al., 2016), the magnitude of the effect suggests that divorce could only explain up to 1.5% of our result if we assume that the children of divorce are 10 percentage points less likely to attend college.

Figure 4 presents cohort-specific estimates that are consistent with the pooled estimates in Table 5. We do not estimate any meaningful differential trends prior to the introduction of the 1970 CAAA, providing further support for the parallel trends assumption.

One might be concerned that the average effects presented in Table 5 are a net zero if income effects result in differential fertility response for men and women. Economic theory suggests that improvements in male labor market conditions should be associated with increases in fertility, while better wages and employment opportunities for women should have opposing income and substitution effects (Schaller, 2016). In Tables B2 and B3 we explore whether there are differential family structure effects by sex. We do not find any effect of particulate matter reductions on the likelihood that women get married, divorced, or on the likelihood of having children, the number of children, or the age at first birth. The estimates are small in magnitude.

In Table B4 we explore whether reductions in particulate matter exposure had effects on who individuals married, an examination of assortative matching. We estimate that a 10 $\mu g/m^3$ reduc-

tion in prenatal TSP exposure is associated with a 63.2 percentage point increase in the likelihood that an individuals' partner was treated, but no meaningful effects on other partner characteristics.

Collectively, these findings suggest that our college attendance effect is unlikely to be driven by differential selection into the second-generation sample, or family structure effects.

4.1.4 Parental Investments

Traditionally, models exploring the mechanisms underlying intergenerational persistence have implicitly assumed that all parents as equally capable. More recently, Cunha and Heckman (2007), Heckman and Mosso (2014), and Becker et al. (2018) model increases in parental human capital as having an effect not only on the earnings of parents but also the production of their children's human capital, introducing a new mechanism — parental investments. If reductions in prenatal pollution exposure increased parental health as well as wealth, then parents may be better placed to spend time and make investments in their children, improving child outcomes (Murnane et al., 2000; Heckman and Carneiro, 2003; Belfield et al., 2006; Cunha et al., 2010; Heckman et al., 2013; Lundberg, 2017; Akee et al., 2018).

One concrete approach to understanding the role that parental investments in cognitive and non-cognitive skills might play is to examine the relationship between prenatal pollution exposure and parental time-use. Table 6 presents estimates of the effect of prenatal pollution exposure on the total amount of quality time spent with children (column 1) and on each of the four categories (column 2-5). We estimate that reductions in prenatal TSP exposure are associated with meaningful increases in quality time spent with children. The reduced form estimates and instrumental variable estimates are similar in magnitude, but the reduced form estimates are more precisely estimated. A 10 $\mu g/m^3$ decrease in prenatal TSP exposure is associated with parents spending an additional 25 minutes of quality time each day. This is a 20% increase relative to the control mean. This increase is largely driven by time spent eating meals together and time spent engaging with the arts.[17]

[17] In appendix Table B5 we estimate significant increases in time spent helping children with homework, time spent eating meals, time spent visiting museums, and time spent watching performing arts. We also estimate meaningful increases in the amount of time spent reading to children, although this effect is not statistically significant at conventional levels.

Figure 5 presents cohort-specific estimates that are consistent with the results in Table 6.We do not estimate any meaningful differential trends prior to the introduction of the 1970 CAAA, providing further support for the parallel trends assumption. We note that the overall effect on quality time (panel a) is more noisily estimated, but that the categories responsible for driving the overall effect in Table 6 are more precisely estimated.

5 Conclusion

We provide early quasi-experimental evidence on the intergenerational consequences of prenatal exposure to ambient air pollution. Exploiting variation in particulate matter induced by the introduction of the 1970 Clean Air Act Amendments — which substantially reduced ambient air pollution — we find that the children of affected cohorts born in nonattainment counties after its implementation were less likely to drop out of high school and more likely to attend college 40-50 years later than those born before. We present evidence to suggest that the intergenerational transmission of this early-life shock to parental endowments arises through increased parental resources and investments.

Back-of-the-envelope calculations suggest that the total first- and second-generation cumulative lifetime earnings effects are 50% larger than the total first-generation cumulative lifetime earnings effects for the 1972 cohort. This combined value is equivalent in magnitude to 55-70% of the mortality benefits estimated in Chay and Greenstone (2003a). To date, the monetized mortality benefits of environmental regulations constitute the vast majority of overall benefits in benefit–cost analysis. We argue that accounting for intergenerational effects would substantially increase the returns to investments in environmental quality. Within-generation estimates of marginal damages substantially underestimate the total welfare effects of improving environmental quality.

It is striking that these effects arise from such short exposure times, early in life. If relatively small changes in air pollution have intergenerational consequences through parental resources and investments, it is likely that larger shocks and policies also have intergenerational consequences.

This reinforces the potential welfare benefits that could be realized through a reallocation of resources from later to earlier in the life cycle. Understanding the mechanisms through which such reallocations can be delivered remains an important area for research.

In turn, our results have important implications for inequality and economic opportunity. It is well established that economic and environmental inequality walk hand-in-hand. Individuals who live in low-income areas and disadvantaged communities are exposed to higher levels of pollution than those that live in high-income areas (Commission for Racial Justice, United Church of Christ, 1987; Mohai et al., 2009; Banzhaf et al., 2019; Colmer et al., 2020; Currie et al., 2020). We document that early-life exposure to pollution has intergenerational consequences that are propagated through economic resources and investments. Our results open up new lines of inquiry, raising questions about the role that investments in environmental quality might play in shaping upward mobility, economic opportunity, inequality and the intergenerational persistence of poverty.

References

Akee, R., W. Copeland, E.J. Costello, and E. Simeonova, "How Does Household Income Affect Child Personality Traits and Behaviors?," *American Economic Review*, 2018, *108* (3), 775–827.

Akresh, R., D. Halim, and M. Kleemans, "Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia," *NBER Working Paper 25265*, 2018.

Almond, D., "Is the 1918 Influenza Pandemic Over? Long-Term Effects of In-Utero Influenza Exposure in the Post-1940 U.S. Population," *Journal of Political Economy*, 2006.

_ , L. Edlund, H. Li, and J. Zhang, "Long-Term Effects of the 1959-61 China Famine: Mainland China and Hong Kong," in T. Ito and A. Rose, eds., *The Economic Consequences of Demographic Change in East Asia*, NBER-EASE Vol. 19, 2010.

Autor, D., D. Dorn, and G. Hanson, "When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men," *American Economic Review: Insights*, 2019, *1* (2), 161–178.

Bailey, M., H. Hoynes, M. Rossin-Slater, and R. Walker, "Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program," *NBER Working Paper 26942*, 2020.

Banzhaf, S., L. Ma, and C. Timmins, "Environmental justice: The economics of race, place, and pollution," *Journal of Economic Perspectives*, 2019, *33* (1), 185–208.

Barr, A. and C. Gibbs, "Breaking the Cycle? The Intergenerational Effects of Head Start," *Mimeo*, 2017.

Bastian, J. and K. Michelmore, "The Long-Term Impact of the Earned Income Tax Credit on Children's Education and Employment Outcomes," *Journal of Labor Economics*, 2018, *36* (4), 1127–1163.

Becker, G., "An Economic Analysis of Fertility," *Demographic and Economic Change in Developed Countries*, 1960.

_ , *A Treatise on the Family*, Harvard University Press, 1981.

_ **and N. Tomes**, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 1979.

_ **, S.D. Kominers, K.M. Murphy, and J.L. Spenkuch**, "A Theory of Intergenerational Mobility," *Journal of Political Economy*, 2018, *126* (S1), S7–S25.

Belfield, C., M. Nores, S. Barnett, and L. Schwienhart, "The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup," *Journal of Human Resources*, 2006.

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

_ **and K. Salvanes**, "Education and Family Background: Mechanisms and Policies," in E.A. Hanushek, S. Machin, and L. Woessmann, eds., *Handbook of the Economics of Education*, Vol. 3, Amsterdam: Elsevier, 2011, pp. 201–247.

_ **and L. Chadwick**, "Intergenerational Income Mobility in Permanent and Separated Families," *Economic Letters*, 2003, *80* (2), 239–246.

_ **, L. Lindahl, and M. Lindquist**, "What More Than Parental Income? An Exploration of What Swedish Siblings Get from Their Parents," *B.E. Journal of Economic Analysis and Policy: Contributions to Economic Analysis and Policy*, 2010, *10* (1).

_ **, M. Jantti, and G. Solon**, "Nature and Nurture in the Intergenerational Transmission of Socioeconomic Status: Evidence from Swedish Children and Their Biological and Rearing Parents," *B.E. Journal of Economic Analysis and Policy: Advances in Economic Analysis and Policy*, 2007, *7* (2).

Black, D., N. Kolesnikova, S.G. Sanders, and L.J. Taylor, "Are Children "Normal"?," *Review of Economics and Statistics*, 2013, *95* (1), 21–33.

Black, S., A. Butikofer, P. Devereux, and K. Salvanes, "This is Only a Test? Long-Run Impacts of Prenatal Exposure to Radioactive Fallout," *Review of Economics and Statistics*, 2018.

_ **and P. Devereux**, "Recent Developments in Intergenerational Mobility," in "Handbook of Labor Economics," Elsevier, 2011.

_ , _ , **and K. Salvanes**, "Why the Apple doesn't fall far: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, 2005.

_ , _ , **and** _ , "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes," *Quarterly Journal of Economics*, 2007.

_ , _ , **P. Lundborg, and K. Majlesi**, "Poor Little Rich Kids? The Role of Nature Versus Nurture in Wealth and Other Economic Outcomes and Behaviors," *Review of Economic Studies*, 2020, *87* (4).

Bleakley, H., "Health, Human Capital, and Development," *Annual Review of Economics*, 2010, *2*, 283–310.

Bové, H., E. Bongaerts, and E. et al. Slenders, "Ambient black carbon particles reach the fetal side of human placenta," *Nature Communications*, 2019, *10* (3866).

Bulman, G., R. Fairlie, S. Goodman, and A. Isen, "Parental Resources and College Attendance: Evidence from Lottery Wins," *Mimeo*, 2017.

Carre, J., N. Gatimel, J. Moreau, J. Parinaud, and R. Leandri, "Does air pollution play a role in infertility? A systematic review," *Environmental Health*, 2017, *16* (82).

Chay, K. and M. Greenstone, "Air Quality, Infant Mortality, and the Clean Air Act of 1970," Working Paper 10053, National Bureau of Economic Research October 2003.

_ **and** _ , "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *The Quarterly Journal of Economics*, 2003, *118* (3), 1121–1167.

_ **and** _ , "Does Air Quality matter? Evidence from the Housing Market," *Journal of Political Economy*, 2005.

Chay, K.Y., C. Dobkin, and M. Greenstone, "The Clean Air Act of 1970 and Adult Mortality," *Journal of Risk and Uncertainty*, 2003.

Chetty, R. and N. Hendren, "The Effects of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *Quarterly Journal of Economics*, 2018.

_ **and** _ , "The Effects of Neighborhoods on Intergenerational Mobility II: County Level Estimates," *Quarterly Journal of Economics*, 2018.

_ **, J. Friedman, N. Hendren, M. Jones, and S. Porter**, "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility," *NBER Working Paper 25147*, 2018.

_ **, N. Hendren, and L. Katz**, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

_ , _ , **M. Jones, and S. Porter**, "Race and Economic Opportunity in the United States: an Intergenerational Perspective," *Quarterly Journal of Economics*, 2019, *135* (2), 711–783.

_ , _ , **P. Kline, and E. Saez**, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Chyn, E., "Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on the Labor Market Outcomes of Children," *American Economic Review*, 2018, *108* (10), 3028–56.

Clay, K., M. Portnykh, and E. Severnini, "Toxic Truth: Lead and Fertility," *NBER Working Paper 24607*, 2018.

Colmer, J., I. Hardman, J. Shimshack, and J. Voorheis, "Disparities in PM2.5 air pollution in the United States," *Science*, 2020, *369* (6503), 575–578.

Commission for Racial Justice, United Church of Christ, "Toxic Wastes and Race in the United States: A National Report on the Racial and Socio-Economic Characteristics of Communities with Hazardous Waste Sites," 1987.

Cunha, F. and J. Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

_ , _ , **and S. Schennach**, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

Currie, J., J. Voorheis, and R. Walker, "What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality," *NBER Working Paper 26659*, 2020.

_ **, M. Neidell, and J.F. Schmieder**, "Air Pollution and Infant Health: Lessons from New Jersey," *Quarterly Journal of Economics*, 2009.

Dejmek, S., G. Selevan, I. Benes, I. Solansky, and R.J. Sram, "Fetal growth and maternal exposure to particulate matter during pregnancy," *Environmental Health Perspectives*, 1999, *107* (6), 475–480.

Deming, D., "The Growing Importance of Social Skills in the Labor Market," *Quarterly Journal of Economics*, 2017, *132* (4), 1593–1640.

Deryugina, T., G. Heutel, H. H. Miller, D. Molitor, and J. Reif, "The mortality and medical costs of air pollution: Evidence from changes in wind direction," *American Economic Review*, 2019, *109* (12).

Dettling, L.J. and M. Kearney, "House Prices and Birth Rates: The Impact of the Real Estate Market on the Decision to have a Baby," *Journal of Public Economics*, 2014, *110*, 82–100.

Deutscher, N., "Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia," *American Economic Journal: Applied Economics*, 2019.

Doepke, M. and F. Zilibotti, "Parenting with Style: Altruism and Paternalism in Inter-generation Preference Transmission," *Econometrica*, 2017, *85* (5), 1331–1371.

_ , **G. Sorrenti, and F. Zilibotti**, "The Economics of Parenting," *Annual Review of Economics*, 2019, *11*, 55–84.

East, C., S. Miller, M. Page, and L. Wherry, "Multi-Generational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health," *NBER Working Paper 23810*, 2017.

Ebenstein, A., V. Lavy, and S. Roth, "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution," *American Economic Journal: Applied Economics*, 2016, *8* (4), 36–65.

Fertig, A.R. and T. Watson, "Minimum Drinking Age Laws and Infant Health Outcomes," *Journal of Health Economics*, 2009.

Frimmel, W., M. Halla, and R. Winter-Ebmer, "How Does Parental Divorce Affect Children's Long-term Outcomes?," *IZA DP No. 9928*, 2016.

Glinianaia, S., J. Ranking, J. Bell, T. Pless-Mulloli, and D. Howel, "Particulate Air Pollution and Fetal Health: A Systematic Review of the Epidemiologic Evidence," *Epidemiology*, 2004, *15* (1), 36–45.

Greenstone, M., "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures," *Journal of Political Economy*, 2002.

_ , **J. List, and C. Syverson**, "The Effects of Environmental Regulation on the Competitiveness of US Manufacturing," *Mimeo*, 2012.

Gronqvist, H., J.P. Nilsson, and P. Robling, "Understanding How Low Levels of Early Lead Exposure Affect Children's Life Trajectories," *Journal of Political Economy*, 2020, *128* (9).

Grossman, M., "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 1972, *80* (2), 223–255.

Gruber, J., "Is Making Divorce Easier Bad for Children? The Long-Run Implications of Unilateral Divorce," *Journal of Labor Economics*, 2004.

Heckman, J. and P. Carneiro, "Human Capital Policy," *NBER Working Paper 9495*, 2003.

_ **and S. Mosso**, "The Economics of Human Development and Social Mobility," *Annual Review of Economics*, 2014, *6* (1), 689–733.

_ **and Y. Rubenstein**, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, 2001, *91* (2), 145–149.

_ , **R. Pinto, and P. Savelyev**, "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, 2013, *103* (6), 2052–2086.

Huifeng, Y., J. Xiaotong, L. Guangke, H. Meng, and S. Nan, "Maternal Exposure to PM2.5 Affects Fetal Lung Development at Sensitive Windows," *Environmental Science & Technology*, 2020, *54* (1), 316–324.

Isen, A., M. Rossin-Slater, and R. Walker, "Every Breath You Take — Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970," *Journal of Political Economy*, 2017.

Kearney, M. and R. Wilson, "Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom," *Review of Economics and Statistics*, 2018, *100* (4), 678–690.

Kelly, E., "The Scourge of Asian Flu: In Utero Exposure to Pandemic Influenza and the Development of a Cohort of British Children," *Journal of Human Resources*, 2011.

Kleven, H., C. Landais, and J.E. Sogaard, "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families," *American Economic Review: Insights*, 2020.

Lee, J. and L. Taylor, "Randomized Safety Inspections and Risk Exposure on the Job: Quasi-experimental Estimates of the Value of a Statistical Life," *American Economic Journal: Applied Economics*, 2019, *11* (4), 350–74.

Lee, S.Y. and A. Seshadri, "On the Intergenerational Transmission of Economic Status," *Journal of Political Economy*, 2019, *127* (2).

Lovenheim, M., "The Effect of Liquid Housing Wealth on College Enrollment," *Journal of Labor Economics*, 2011, *29* (4), 741–771.

_ **and C. Lockwood Reynolds**, "The Effect of Housing Wealth on College Choice: Evidence from the Housing Boom," *Journal of Human Resources*, 2013, *14* (1), 3–37.

_ **and K.J. Mumford**, "Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market," *Review of Economics and Statistics*, 2013, *95* (2), 464–475.

Lundberg, S., "Non-Cognitive Skills as Human Capital," *Mimeo*, 2017.

Mohai, P., D. Pellow, and J.T. Roberts, "Environmental Justice," *Annual Review of Environment and Resources*, 2009, *34*, 405–430.

Murnane, R., J. Willett, Y. Duhaldeborde, and J. Tyler, "How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?," *Journal of Policy Analysis and Management*, 2000, *19* (4), 547–568.

Nilsson, J.P., "Alcohol Availability, Prenatal Conditions, and Long-term Economic Outcomes," *Journal of Political Economy*, 2017, *125* (4).

Painter, G. and D.I. Levine, "Family Structure and Youths' Outcomes: Which Correlations are Causal?," *Journal of Human Resources*, 2000, *35* (3), 524–549.

Price, J., "Parent-Child Quality Time: Does Birth Order Matter?," *Journal of Human Resources*, 2008, *43* (1), 240–265.

Sanders, N.J. and C. Stoecker, "Where have all the young men gone? Using Sex Ratios to Measure Fetal Death Rates," *Journal of Health Economics*, 2015.

Schaller, J., "Booms, Busts, and Fertility," *Journal of Human Resources*, 2016, *51* (1), 1–29.

Solon, G., "Intergenerational Income Mobility in the United States," *American Economic Review*, 1992.

Voorheis, J., "Air Quality, Human Capital Formation and the Long-term Effects of Environmental Inequality at Birth," Working Paper 2017-05, US Census Bureau 2017.

Wagner, D. and M. Layne, "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications (CARRA) Record Linkage Software," *Mimeo*, 2014.

Walker, R., "Environmental Regulation and Labor Reallocation," *American Economic Review: Papers and Proceedings*, 2011.

_ , "The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce," *Quarterly Journal of Economics*, 2013.

Weinberger, C., "The Increasing Complementarity between Cognitive and Social Skills," *Review of Economics and Statistics*, 2014, *96* (5), 849–861.

Zivin, J. Graff and M. Neidell, "Environment, Health, and Human Capital," *Journal of Economic Literature*, 2013, *51* (3), 689–730.

Tables and Figures

Table 1: The Relationship between Nonattainment and Pollution Exposure Over the Life Cycle

	(1) First-Gen Prenatal Exposure ($\mu g/m^3$)	(2) First-Gen Later-Life Exposure ($\mu g/m^3$)	(3) Second-Gen Prenatal Exposure ($\mu g/m^3$)	(4) Second-Gen Later-Life Exposure ($\mu g/m^3$)
Nonattainment × Post	-11.900***	-0.173	0.688	0.192
	(2.893)	(0.484)	(0.405)	(0.531)
First-Gen Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month			
First-Gen Prenatal Controls	Yes	Yes	Yes	Yes
Observations	122,000	65,500	95,000	61,500
Control Mean	91.99	46.49	57.49	41.75

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. Source: Census Numident, Decennial Census 2000 and 2010 and ACS 2005 through 2015. All regressions include individual demographic controls including sex, race/ethnicity and quadratic in age, pre-1970 CAAA county of birth economic characteristics (employment, total transfer income, personal income per capita) interacted with quadratic trends, county of birth temperature and precipitation measured during the individual's nine month gestational period. Standard errors are clustered at the first generation's county of birth level.

Table 2: The Relationship between First-Generation Prenatal TSP Exposure and Second-Generation College Attendance

	Second-Generation College Attendance (1)	(2)	(3)
Panel A: IV			
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.0227^{**}	-0.0251^{**}	-0.0233^{*}
	(0.0105)	(0.0124)	(0.0124)
Panel B: Reduced Form			
Nonattainment × Post	0.0292^{**}	0.0298^{**}	0.0272^{**}
	(0.0126)	(0.0130)	(0.0129)
First-Gen Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month		
First-Gen Controls	No	Yes	Yes
Second-Gen Controls	No	No	Yes
Second-Gen Fixed Effects?	No	No	Yes
Observations	122,000	122,000	121,000
Control Mean	0.438	0.438	0.438
First Stage F-Stat	25.12	16.93	17.52

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

Table 3: The Relationship between First-Generation Prenatal TSP Exposure and Second-Generation College Attendance: Adopted and Step-Children vs. Biological Children

	Adopted or Step- Child (1)	Second-Generation College Attendance (2)	(3)
Panel A: IV			
First-Gen Prenatal TSP ($10\mu g/m^3$)	0.0125	-0.0231*	-0.0472**
	(0.0098)	(0.0137)	(0.0232)
Panel B: Reduced Form			
Nonattainment × Post	-0.0148	0.0272*	0.0574**
	(0.0115)	(0.0151)	(0.0248)
Child Type	–	Biological Child	Adopted/Step-Child
First-Gen Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month		
First-Gen Controls	Yes	Yes	Yes
Observations	122,000	89,500	33,000
Control Mean	0.257	0.439	0.437
First Stage F-Stat	16.93	15.85	17.32

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

Table 4: The Relationship between First-Generation Prenatal TSP Exposure and First-Generation Labor Market Outcomes

	(1) logEARNINGS	(2) LABOR FORCE PARTICIPATION	(3) UNEMPLOYED	(4) PUBLIC ASSISTANCE
Panel A: IV				
First-Gen Prenatal TSP ($10\mu g/m^3$)	–0.00668*	-0.00138	–0.00024	-0.00005
	(0.00398)	(0.00088)	(0.00083)	(0.00049)
Panel B: Reduced Form				
Nonattainment × Post	0.00905*	0.001864	0.000323	0.000070
	(0.00533)	(0.00114)	(0.00112)	(0.00066)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month			
Individual Controls	YES	YES	YES	YES
County-level Controls	YES	YES	YES	YES
Observations	2,015,000	2,015,000	2,015,000	2,015,000
Control Mean	$40,600	0.955	0.0353	0.0149
First Stage F-Stat	23.81	23.81	23.81	23.81

NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

Table 5: The Relationship between First-Generation Prenatal TSP Exposure and First-Generation Family Structure Outcomes

	(1) ANY KIDS	(2) # OF CHILDREN	(3) AGE AT FIRST BIRTH	(4) EVER MARRIED	(5) DIVORCED
Panel A: IV					
First-Gen Prenatal TSP ($10\mu g/m^3$)	0.000790	0.00734	-0.00367	0.00463**	0.00311*
	(0.0148)	(0.0773)	(0.0272)	(0.00181)	(0.00177)
Panel B: Reduced Form					
Nonattainment × Post	-0.001097	-0.0102	0.004971	-0.006260**	-0.004199*
	(0.00204)	(0.01082)	(0.03663)	(0.00209)	(0.00218)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month				
Individual Controls	YES	YES	YES	YES	YES
County-level Controls	YES	YES	YES	YES	YES
Observations	2,699,000	2,699,000	2,015,000	2,015,000	1,703,000
Control Mean	0.802	1.758	25.68	0.868	0.135
First Stage F-Stat	20.91	20.91	23.81	23.81	23.32

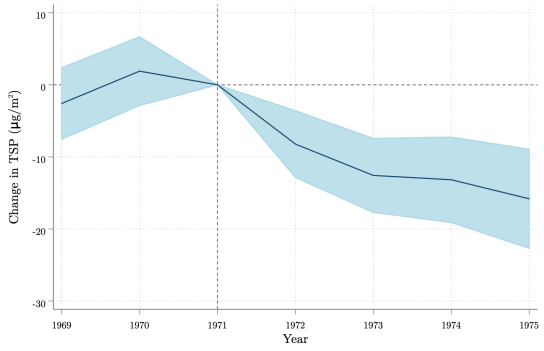
NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

Table 6: The Relationship between First-Generation Prenatal TSP Exposure and Parent-Child Interactions

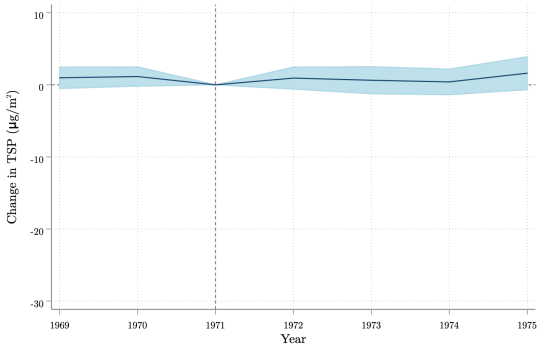
	(1) Quality Time (minutes)	(2) Helping Children (minutes)	(3) Eating Meals Together (minutes)	(4) Arts and Sports (minutes)	(5) Child Care (minutes)
Panel A: IV					
First-Gen Prenatal TSP ($10\mu g/m^3$)	-25.280*	-5.092	-10.960	-10.240*	1.008
	(14.430)	(8.214)	(7.677)	(5.293)	(7.471)
Panel B: Reduced Form					
Nonattainment × Post	24.250**	4.883	10.510*	9.818**	-0.967
	(12.320)	(7.968)	(6.065)	(3.896)	(7.111)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month and Interview Day				
Individual Controls	Yes	Yes	Yes	Yes	Yes
County-level Controls	Yes	Yes	Yes	Yes	Yes
Observations	4,800	4,800	4,800	4,800	4,800
Control Mean	126.3	34.37	42.54	7.8	41.61
First Stage F-Stat	8.659	8.659	8.659	8.659	8.659

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

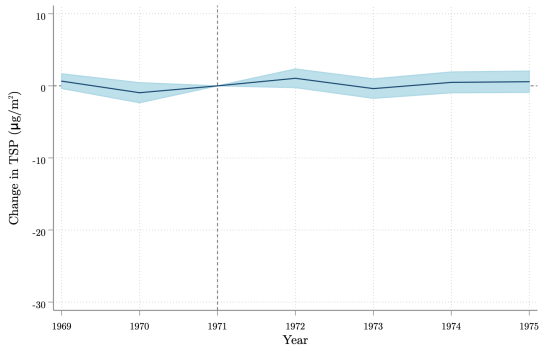
Figure 1: The Relationship between First-Generation Prenatal Nonattainment Exposure and TSP Exposure



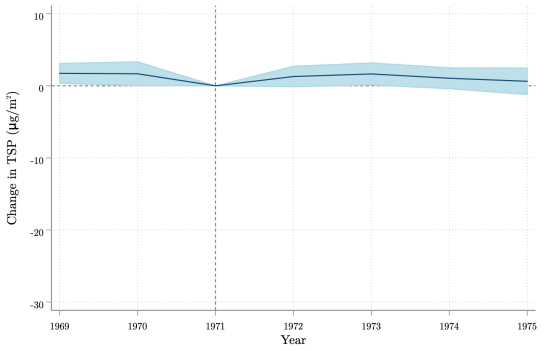
(a) First-Gen Prenatal TSP Exposure



(b) First-Gen Adult TSP Exposure



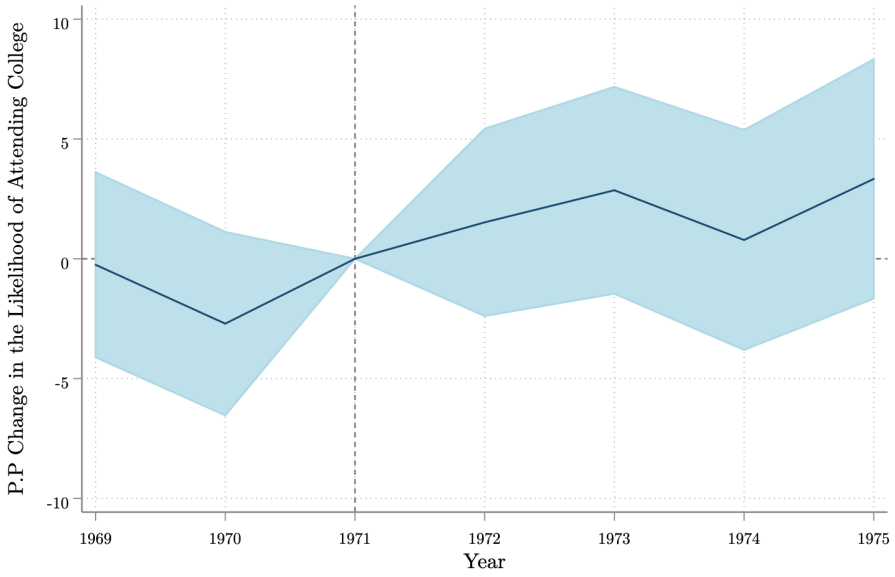
(c) Second-Gen Prenatal TSP Exposure



(d) Second-Gen Pre-College TSP Exposure

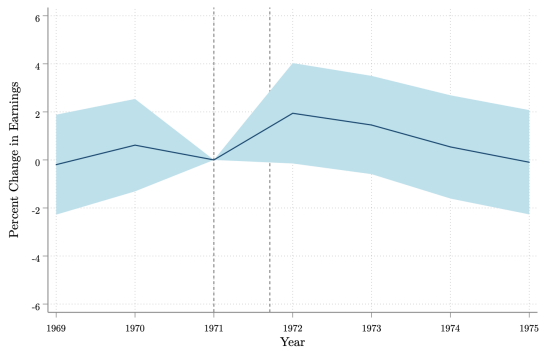
Notes: These figures present cohort-specific estimates of the association between first-generation prenatal exposure to nonattainment and TSP exposure. Panel a) presents estimates of the association between first-generation prenatal exposure to nonattainment and first-generation prenatal TSP exposure. This is the first-stage of our analysis. Panel b) presents estimates of the association between first-generation prenatal exposure to nonattainment and first-generation TSP exposure in adulthood. Panel c) presents estimates of the association between first-generation prenatal exposure to nonattainment and second-generation TSP exposure during gestation. Panel d) presents estimates of the association between first-generation prenatal exposure to nonattainment and second-generation TSP exposure during high school. We include first-generation county-of-birth, state-of-birth $\times$ year, and birth month fixed effects. No other controls are included. Results are robust to the inclusion of first-generation prenatal controls and to the inclusion of second-generation fixed effects and controls. See Table 1 for more information.

Figure 2: The Relationship between First-Generation Prenatal Nonattainment Exposure and Second-Generation College Attendance

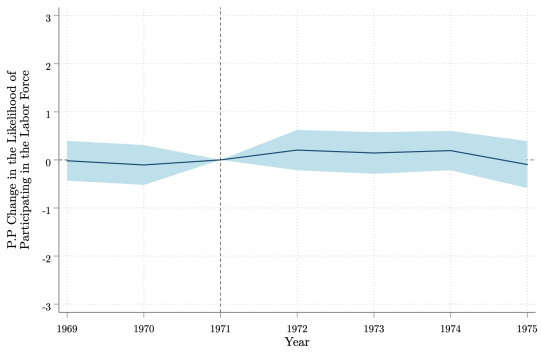


Notes: This figure presents cohort-specific estimates of the association between first-generation prenatal exposure to nonattainment and second-generation college attendance. We include first-generation county-of-birth, state-of-birth × year, and birth month fixed effects (analogous to Table 2, column 1). No other controls are included. Results are robust to including first-generation prenatal controls (analogous to Table 2, column 2)and to including second-generation fixed effects and controls (analogous to Table 2, column 3). See Table 1 for more information.

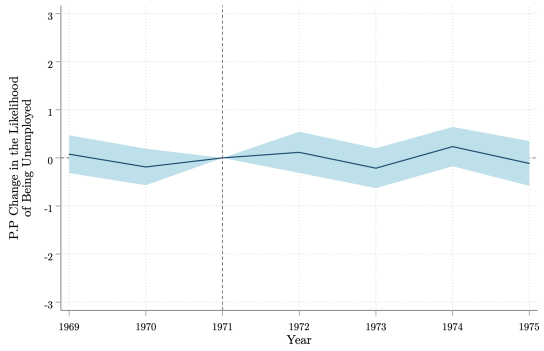
Figure 3: The Relationship between First-Generation Prenatal Nonattainment Exposure and First-Generation Economic Outcomes



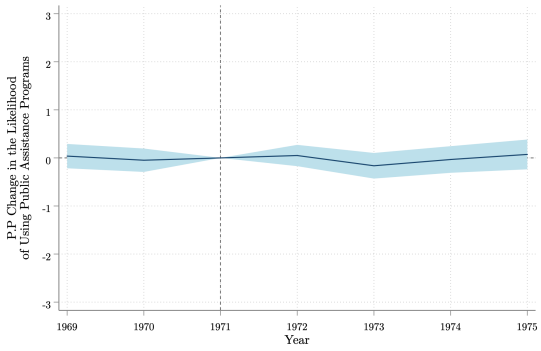
(a) Earnings



(b) Labor Force Participation



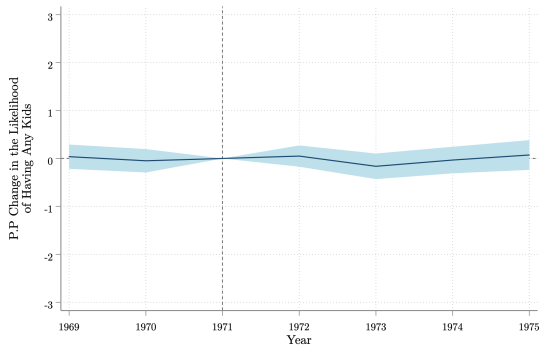
(c) Unemployed



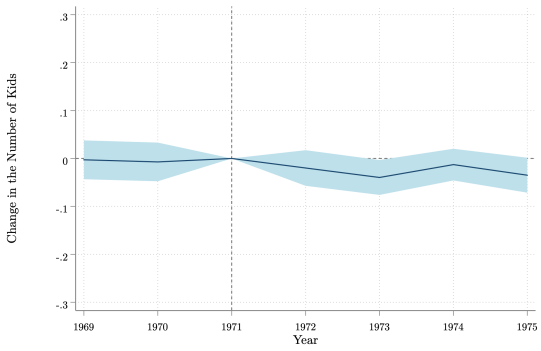
(d) Public Assistance

Notes: These figures present cohort-specific estimates of the association between first-generation prenatal exposure to nonattainment and first-generation economic outcomes. Panel a) presents estimates of the association between first-generation prenatal exposure to nonattainment and later-life earnings. Panel b) presents estimates of the association between first-generation prenatal exposure to nonattainment and the likelihood of being in the labor force. Panel c) presents estimates of the association between first-generation prenatal exposure to nonattainment and the likelihood of being unemployed. Panel d) presents estimates of the association between first-generation prenatal exposure to nonattainment and the likelihood of requiring public assistance. We include first-generation county-of-birth, state-of-birth $\times$ year, and birth month fixed effects. No other controls are included. Results are robust to the including individual-level and county-level controls as presented in Table 4. See Table 1 for more information.

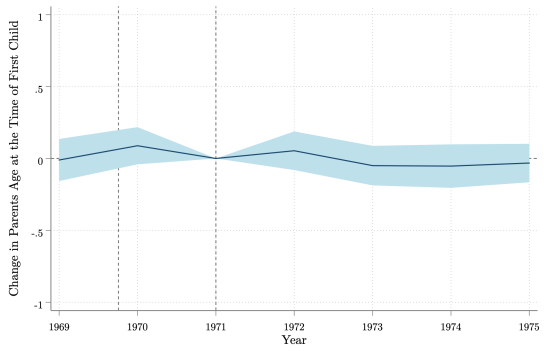
Figure 4: The Relationship between First-Generation Prenatal Nonattainment Exposure and First-Generation Family Structure



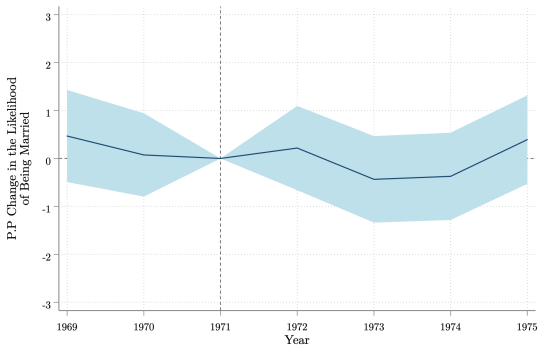
(a) Any Kids



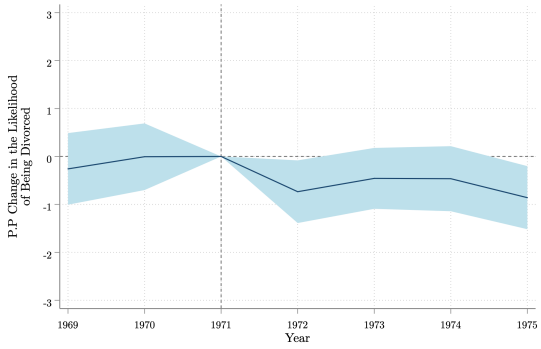
(b) Number of Kids



(c) Age at Time of First Child



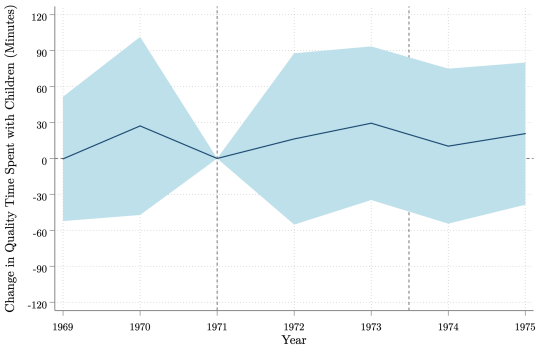
(d) Ever Married



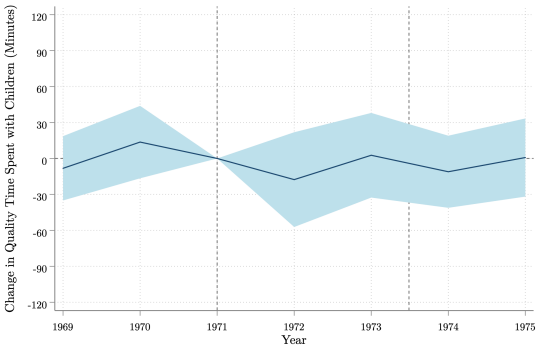
(e) Ever Divorced

Notes: These figures present cohort-specific estimates of the association between first-generation prenatal exposure to nonattainment and first-generation family structure outcomes. Panel a) presents estimates of the association between first-generation prenatal exposure to nonattainment and the likelihood of having any kids. Panel b) presents estimates of the association between first-generation prenatal exposure to nonattainment and the number of kids. Panel c) presents estimates of the association between first-generation prenatal exposure to nonattainment and age at the time of the first child was born. Panel d) presents estimates of the association between first-generation prenatal exposure to nonattainment and the likelihood of ever having been married. Panel e) presents estimates of the association between first-generation prenatal exposure to nonattainment and the likelihood of being divorced. We include first-generation county-of-birth, state-of-birth $\times$ year, and birth month fixed effects. No other controls are included. Results are robust to the including individual-level and county-level controls as presented in Table 5. See Table 1 for more information.

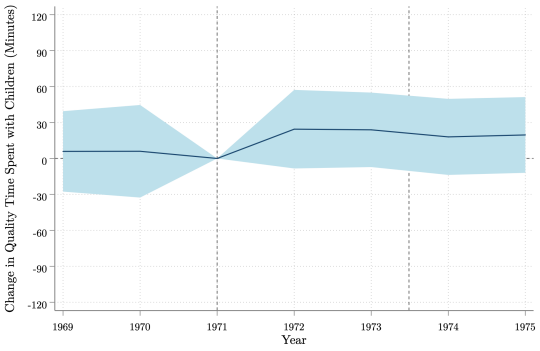
Figure 5: The Relationship between First-Generation Prenatal Nonattainment Exposure and Parent-Child Interactions



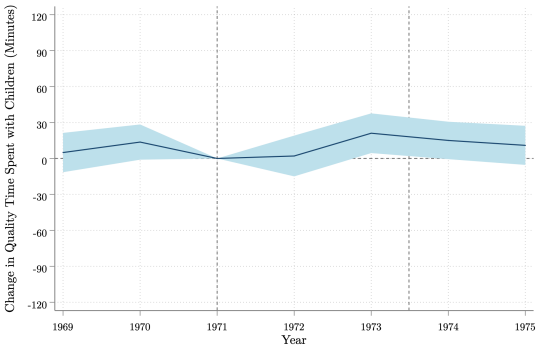
(a) Total Quality Time



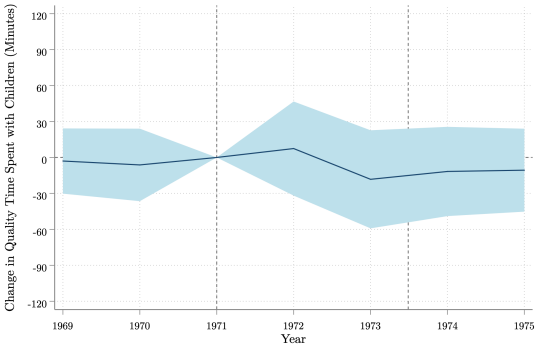
(b) Time Spent Helping Children



(c) Time Spent Eating Meals Together



(d) Time Spent on Arts and Sports



(e) Time Spent on Child Care

Notes: These figures present cohort-specific estimates of the association between first-generation prenatal exposure to nonattainment and parent-child interactions. Panel a) presents estimates of the association between first-generation prenatal exposure to nonattainment and the total amount of time quality time spent with children. Panel b) presents estimates of the association between first-generation prenatal exposure to nonattainment and the amount of time spent helping children. Panel c) presents estimates of the association between first-generation prenatal exposure to nonattainment and the total amount of time spent eating meals together. Panel d) presents estimates of the association between first-generation prenatal exposure to nonattainment and the total amount of time spent on arts and sports. Panel e) presents estimates of the association between first-generation prenatal exposure to nonattainment and the total amount of time spent looking after children as the primary activity. We include first-generation county-of-birth, state-of-birth $\times$ year, and birth month fixed effects. No other controls are included. Results are robust to the including individual-level and county-level controls as presented in Table 6. See Table 1 for more information.

Appendix Materials— For Online Publication

Contents

A Main Results: Additional Results and Robustness Tests

In this appendix we present a series of additional results and robustness tests supporting our main result, documenting the effects of first-generation prenatal TSP exposure on second-generation college attendance.

A.1 Descriptive Statistics

Table A1: Descriptive Statistics Prior to the 1970 CAAA

	(1) Nonattainment County	(2) Attainment County	(3) Difference (2)-(1)
Cohort Demographics			
Female	0.430	0.425	0.005**
	(0.00157)	(0.00229)	(0.00278)
Black	0.139	0.105	0.034
	(0.0129)	(0.0174)	(0.0216)
White	0.671	0.694	-0.023
	(0.0185)	(0.0271)	(0.0328)
Hispanic	0.094	0.082	0.012
	(0.0181)	(0.0219)	(0.0284)
Other	0.010	0.023	-0.013*
	(0.0012)	(0.0067)	(0.0068)
Environmental Exposure			
Gestational Exposure (10 $\mu g/m^3$)	11.83	7.93	3.895***
	(0.3172)	(0.2464)	(0.4017)
Rainfall (days)	23	25.5	-2.505
	(1.22)	(1.17)	(1.69)
Avg. Daily Maximum Temperature	17.95	18.33	-0.3831
	(0.5595)	(0.5903)	(0.8133)
County Socio-Economic Characteristics			
County Population	1,628,000	458,000	1,170,000**
	(512,000)	(54,740)	(514,900)
Personal Income per Capita ($ Year)	4,245	4,003	242*
	(113.9)	(92.51)	(146.8)

NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. Standard errors are clusterd at the parent's county of birth level. Source: Census Numident, Decennial Census 2000 and 2010, ACS 2005 through 2015.

A.2 Migration

Table A2: The Effect of First-Generation Prenatal TSP Exposure on Migration out of First-gen County of Birth

	(1) First-Gen Migration (Out of County)	(2) Second-Gen Migration (Out of County)
Panel A: IV		
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.00403	0.00891
	(0.0109)	(0.0101)
Panel B: Reduced Form		
Nonattainment × Post	0.00477	-0.01062
	(0.0128)	(0.0120)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month	
First-Gen Controls	Yes	Yes
Observations	119,000	121,000
Control Mean	0.560	0.480
First Stage F-Stat	21.68	21.68

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. Source: Census Numident, Decennial Census Short Form 2000 and 2010, ACS 2005 through 2015 and the Census MAFARF. See Table 1 for more information.

A.2.1 Maternal vs. Paternal Exposure

In this section we explore the effects of first-generation TSP exposure on second-generation college attendance restricting the treatment group to: cases in which mother's were affected; cases in which father's were affected; cases in which both the mother and father were affected. We estimate slightly larger effects when both the mother and father were affected.

Table A3: The Effect of First-Generation Prenatal TSP Exposure on Second-Generation College Attendance (Maternal vs. Paternal Exposure)

	SECOND-GENERATION COLLEGE ATTENDANCE		
	Mother Exposed	Father Exposed Exposed	Both Exposed
	(1)	(2)	(3)
Panel A: IV			
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.02567*	-0.03086**	-0.0375**
	(0.01321)	(0.01423)	(0.01618)
Panel B: Reduced Form			
Nonattainment × Post	0.03061**	0.03584**	0.0442***
	(0.01397)	(0.01453)	(0.01638)
First-Gen Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month		
First-Gen Controls	NO	YES	YES
Second-Gen Controls	NO	NO	YES
Second-Gen Fixed Effects?	NO	NO	YES
Observations	110,000	99,000	86,500
Control Mean	0.442	0.442	0.442
First Stage F-Stat	16.27	19.97	20.08

NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

A.3 Alternate Sample Restrictions

In our main analysis we focus on a sample that is based on certain parent–child links, include all births between 1969 and 1980, and define "treatment" to be assigned if any parent–child link was affected. The following results show that our findings are robust to using a more narrow window of births around the 1970 CAAA, to including probabilistic parent–child links, and to restricting "treatment definitions" based on maternal or paternal exposure.

A.3.1 Different Treatment Windows

Here we present estimates of the effects of first-generation TSP exposure on second-generation college attendance, restricting the sample to a narrower treatment window: 1969-1975.

Table A4: The Effect of First-Generation Prenatal TSP Exposure on Second-Generation College Attendance (1969-1975 Window)

	SECOND-GENERATION COLLEGE ATTENDANCE		
	(1)	(2)	(3)
Panel A: IV			
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.02783**	-0.02545*	-0.02543*
	(0.01266)	(0.01399)	(0.01445)
Panel B: Reduced Form			
Nonattainment × Post	0.03165**	0.02786**	0.02686**
	(0.01282)	(0.01346)	(0.01337)
First-Gen Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month		
First-Gen Controls	NO	YES	YES
Second-Gen Controls	NO	NO	YES
Second-Gen Fixed Effects?	NO	NO	YES
Observations	110,000	110,000	109,000
Control Mean	0.441	0.441	0.441
First Stage F-Stat	21.55	14.97	15.5

NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

A.3.2 Including Probabilistic Links

Here we present results on the effects of first-generation TSP exposure on second-generation college attendance, using an expanded sample that incorporates probabilistic parent–child links. This increases our sample size from 122,000 to 193,000. We estimate similar effects, although the magnitudes are attenuated consistent with the introduction of measurement error that arises from introducing false links.

Table A5: The Effect of First-Generation Prenatal TSP Exposure on Second-Generation College Attendance (Probabilistic Links)

	SECOND-GENERATION COLLEGE ATTENDANCE		
	(1)	(2)	(3)
Panel A: IV			
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.0147*	-0.0157	-0.0147
	(0.00838)	(0.00970)	(0.00977)
Panel B: Reduced Form			
Nonattainment × Post	0.01906*	0.01894*	0.01752
	(0.01033)	(0.01082)	(0.01085)
First-Gen Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month		
First-Gen Controls	NO	YES	YES
Second-Gen Controls	NO	NO	YES
Second-Gen Fixed Effects?	NO	NO	YES
Observations	193,000	193,000	190,000
Control Mean	0.448	0.448	0.448
First Stage F-Stat	30.23	20.36	20.74

NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

A.4 The Effect of First-Generation Prenatal TSP Exposure on the likelihood of Second-Generation High School Completion

Table A6: The Effect of First-Generation Prenatal TSP Exposure on Second-Generation High School Completion

	SECOND-GENERATION HIGH SCHOOL DROPOUT		
	(1)	(2)	(3)
Panel A: IV			
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.00841	-0.00734	-0.00149
	(0.00532)	(0.00590)	(0.00591)
Panel B: Reduced Form			
Nonattainment × Post	0.0108	0.008747	0.001739
	(0.006582)	(0.006797)	(0.006898)
First-Gen Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month		
First-Gen Controls	NO	YES	YES
Second-Gen Controls	NO	NO	YES
Second-Gen Fixed Effects?	NO	NO	YES
Observations	122,000	122,000	121,000
Control Mean	0.109	0.109	0.109
First Stage F-Stat	25.12	16.93	17.52

NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

B Exploring Mechanisms: Additional Results and Robustness Tests

B.1 First-Generation Educational Outcomes

Table B1: The Effect of First-Generation Prenatal TSP Exposure on First-Generation Educational Outcomes

	(1) Years of Schooling	(2) Attended College	(3) High School Dropout
Panel A: IV			
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.0097	0.0008	0.0005
	(0.00115)	(0.00023)	(0.00011)
Panel B: Reduced Form			
Nonattainment × Post	0.01319	-0.00111	-0.0007
	(0.01497)	(0.00316)	(0.00141)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month		
Individual Controls	Yes	Yes	Yes
County-level Controls	Yes	Yes	Yes
Observations	2,015,000	2,015,000	2,015,000
Control Mean	14.76	0.701	0.0436
First Stage F-Stat	23.81	23.81	23.81

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

B.2 Family Structure Results by Sex

Table B2: The Effect of First-Generation Prenatal TSP Exposure on First-Generation Family Structure Outcomes (Women)

	(1) Any Kids	(2) # of Children	(3) Age at First Birth	(4) Ever Married	(5) Divorced
Panel A: IV					
First-Gen Prenatal TSP ($10\mu g/m^3$)	0.00128	0.01582	0.00043	0.00025	0.00006
	(0.00186)	(0.01022)	(0.00362)	(0.00028)	(0.00020)
Panel B: Reduced Form					
Nonattainment × Post	-0.00178	-0.02202	-0.00589	-0.00338	-0.00084
	(0.00257)	(0.01418)	(0.04933)	(0.00369)	(0.00277)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month				
Individual Controls	Yes	Yes	Yes	Yes	Yes
County-level Controls	Yes	Yes	Yes	Yes	Yes
Observations	1,370,000	1,370,000	910,000	910,000	910,000
Control Mean	0.8308	1.871	26.57	0.8062	0.095
First Stage F-Stat	20.6	20.6	22.18	22.18	22.18

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

Table B3: The Effect of First-Generation Prenatal TSP Exposure on First-Generation Family Structure Outcomes (Men)

	(1) Any Kids	(2) # of Children	(3) Age at First Birth	(4) Ever Married	(5) Divorced
Panel A: IV					
First-Gen Prenatal TSP ($10\mu g/m^3$)	0.000291	-0.001274	-0.00482	0.00045	0.00552**
	(0.01906)	(0.1048)	(0.03769)	(0.00265)	(0.00227)
Panel B: Reduced Form					
Nonattainment × Post	-0.00040	0.00177	0.00649	-0.00061	-0.00742***
	(0.00264)	(0.01449)	(0.05052)	(0.00358)	(0.00258)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month				
Individual Controls	Yes	Yes	Yes	Yes	Yes
County-level Controls	Yes	Yes	Yes	Yes	Yes
Observations	1,329,000	1,329,000	1,105,000	1,105,000	1,105,000
Control Mean	0.772	1.641	24.94	0.708	0.133
First Stage F-Stat	21.24	21.24	25.15	25.15	25.15

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

B.3 Assortative Matching

Table B4: The Effect of First-Generation on Partner Characteristics

	(1) Partner Treated	(2) Both Employed	(3) Both College Educated	(4) Same Income Quintile
Panel A: IV				
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.632***	0.0027	-0.0040	-0.0054
	(0.121)	(0.0023)	(0.0028)	(0.0035)
Panel B: Reduced Form				
Nonattainment × Post	0.859***	-0.00364	0.00543	0.00739
	(0.00379)	(0.00301)	(0.00357)	(0.00451)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month			
First-Gen Controls	YES	YES	YES	YES
Observations	812,000	812,000	812,000	812,000
Control Mean	0	0.864	0.286	0.286
First Stage F-Stat	27.16	27.16	27.16	27.16

NOTES: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. See Table 1 for more information.

B.4 Detailed Time-Use Results

Table B5: The Effect of In-Utero TSP Exposure on Parental Time-Use (Detailed Categories)

	(1) Art (minutes)	(2) Meals (minutes)	(3) Teaching (minutes)	(4) Homework (minutes)	(5) Child Care (minutes)	(6) Museums (minutes)	(7) Playing ,not sports (minutes)	(8) Reading (minutes)	(9) Religious activities (minutes)	(10) Playing , sports (minutes)	(11) Talking (minutes)	(12) Performing arts (minutes)	(13) Physical care (minutes)
Panel A: IV													
First-Gen Prenatal TSP ($10\mu g/m^3$)	-0.1654	-10.96	-0.1413	-2.154*	1.616	-3.016*	-1.487	-2.008	-4.789	-0.4274	0.8642	-2.005	-0.6077
	(0.49)	(7.677)	(0.202)	(1.253)	(3.602)	(1.789)	(7.466)	(1.405)	(3.70)	(0.9547)	(1.532)	(1.351)	(6.582)
Panel B: Reduced Form													
Nonattainment × Post	0.1586	10.51*	0.1355	2.066*	-1.55	2.892**	1.426	1.926	4.593	0.4099	-0.8288	1.923*	0.5828
	(0.4746)	(6.065)	(0.1878)	(1.062)	(3.312)	(1.416)	(7.218)	(1.201)	(3.135)	(0.929)	(1.458)	(1.127)	(6.32)
Fixed Effects	County-of-birth, State-of-birth × Year, Birth Month and Interview Day												
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-level Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800
Control Mean	0.5685	42.54	0.0765	1.946	6.71	0.7263	25.85	2.986	5.966	0.7527	2.941	0.3553	34.9
First Stage F-Stat	8.659	8.659	8.659	8.659	8.659	8.659	8.659	8.659	8.659	8.659	8.659	8.659	8.659

Notes: Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. Source: Decennial Census Short Form 2000 and 2010, ATUS 2003-2017. See Table 1 for more information.

C Data Appendix

C.1 The Census Bureau's Data Linkage Infrastructure

Note: this section describes the overall Census data linkage infrastructure. The procedures described here have already been performed on the analysis data we work with. This appendix describes the use of Census, survey and administrative data to identify parent–child links.

The U.S. Census Bureau is authorized, under Titles 13 and 26 of the US Code, to utilize all available data resources, including administrative records and commercially provided data, to improve the measurement of the US population and economy. Under this authority, the Census Bureau has developed a data linkage infrastructure which allows researchers to integrate data from multiple sources, including administrative records from federal and state government agencies, Decennial Census data, and demographic surveys. The central component of the Data linkage infrastructure is the Person Identification Validation System (PVS), which is described in further detail in Wagner and Layne (2014).

PVS is designed as a flexible probabilistic matching system that can be deployed in production to analyze very large datasets in a computationally efficient manner. PVS has two components: a person-based matching algorithm and an address-based matching algorithm. The address based matching algorithm takes a string address as an input (e.g. "1600 Pennsylvania Ave NW, Washington, DC 20001"), splits the string into components (street number, street name, street suffix, city, state, zipcode), standardizes these components, and then matches the address to a reference file (the Census Master Address File), optimizing on a fuzzy string comparator (the Levanstein string distance). The person based matching algorithm has a similar structure: it takes as input the available personally identifiable information on a file (name, SSN, date of birth, sex, address), and, after standardization, matches these PII fields to a separate reference file (the Census Numident).

Each of these matching algorithms produces a unique anonymized identifier for each successful match. For the address matching algorithm, the resulting identifier is called a MAFID (Master Address File Identifier), while the person-based matching algorithm uses PIKs (protected identification keys). MAFIDs and PIKs are both static hashes referencing a single entity in the relevant reference file, and can thus be used to link datasets without including any personally identifiable information on the research files used by researchers. Any attempt to infer PII from a research file with PIKs or MAFIDs is thus a violation of Title 13, with potential punishments including 10 years in prison, and hundreds of thousands of dollars in fines.

Not all of the PII inputs used by the PVS system are found in every microdata file on which PVS is applied. In particular, Social Security numbers are rarely elicited on demographic surveys, and have never been asked for in decennial Censuses.[18] Administrative records which contain SSNs (e.g. most tax records) can be assigned a PIK in 99+ percent of cases. Match rates are still high for many demographic surveys and the decennial census, which ask for name and exact date of birth. The PIK assignment rate for the 2010 Census is about 91 percent, while the PIK assignment rate for the 2013 American Community Survey is about 94 percent.

[18]The Current Population Survey ASEC asked for SSNs until 2002; however, non-response increased dramatically through the 1990s. This was in fact one of the motivating factors in the development of PVS. Moving from SSN-based matching to PVS-based probabilistic matching actually increased match rates for the CPS after 2002.

C.2 Parent–Child Links

To study the intergenerational effects of the Clean Air Act, it is necessary to locate parents at birth (around the enactment of 1970 CAAA), link these parents to their children, and measure outcomes for both parents and children. We begin by assembling a database of all parent–child links that can be evaluated using the various data sources available in the Census Data Linkage Infrastructure. The set of links we are able to identify is not, we should stress, the full population of links. In our empirical analysis, we will attempt to re-weight the data to address the fact that the missing links we are not able to identify are almost certainly not missing at random.

To benchmark our link coverage, consider that the completed cohort fertility rate for women born in 1970 is about 2.1. There were about 44 million women aged 30-50 in the 2010 Census (i.e. born between 1960-1980). Taking the 1970 CCFR as constant throughout this group, we can expect at most 92 million biological children. In practice we will identify fewer than this, due to linkage error, and the fact that women born in the latter part of our birth year range will not have completed fertility in the latest available data we are using to identify parent–child links (the 2015 ACS).

C.2.1 Decennial Census Data

The 2000 and 2010 decennial Census 100 percent detail file (HDF), colloquially the "Census short form", collects an abbreviated set of demographic information from the full population of the United States in decadal Census years. This demographic information includes date of birth, sex, race and ethnicity, and some relationship information. Unfortunately, the relationship information collected in the Census does not capture the full relationship structure within a household. Rather, the Census collects information from each individual in a household on their relationship to the primary household member (the first person listed on the census form for the household), coded as the variable QREL.

This means it is possible to identify two types of parent child links:"certain" parent–child links between a child and the householder parent, and "probable" parent–child links between a child and the married or unmarried partner of their parent householder. The relationship codes are sufficiently detailed to separate biological children of a householder (QREL code 3), adopted children (QREL code 4) and stepchildren (QREL code 5). For the purposes of the project at hand, we identify only parent–child links (certain or probable) for parents born between 1960-1980.

To identify these two types of links in the 2000 Census HDF, we use the following algorithm. We first subset the HDF by age and relationship code, retaining only individuals aged 40 or younger (i.e. who were born after 1960) who have QREL codes 1 (householder), 2 (spouse of householder), 3 (biological child of householder), 4 (adopted child of householder, 5 (stepchild of householder) or 19 (unmarried partner of householder). Then, for each household, we assign three link variables: "Certain Parent", which is the PIK of the householder, "Probable-Married", which is the PIK of the householder's spouse, and "Probable-Unmarried", which is the PIK of the householder's unmarried partner. Each of these variables are missing if the relevant PIK is missing (due to PIK non-assignment when the HDF was analyzed via PVS). We then reshape the data into long form (so each row contains the child's PIK, the parent's PIK and indicators for the type of child and the type of parent). We discard all cases where the child or parent's PIKs are missing.

This yields a dataset containing about 65 million parent–child links. Of these, about 35 million

are "Certain" Links", about 28 million are "Probable-Married", and the remaining approximately 2 million are "Probable-Unmarried". We identify more mother-child links ($\approx$ 38 million) than father-child links (27 million). As expected, the parent–child links identified in the 2000 HDF are heavily tilted toward the older parents: about 51 million links involve parents born before 1970, while about 14 million involve parents born after 1980.

We repeat the use of the same algorithm to identify parent–child links in the 2010 HDF. We identify substantially more links in the 2010 Census, as expected. In all, we identify 115 million parent–child links—of these, about 64 million are "Certain", 46 million are "Probable-Married" and the remaining 5 million are "Probable-Unmarried". As with the 2000 HDF, we identify more mother-child links (65 million) compared to father-child links (50 million). We continue to identify more parent–child links for parents born before 1970, although the split is much more even compared to the 2000 HDF (reflecting the fact that women born before 1970 had largely completed fertility, while women born after 1970 were still in prime childbearing age ranges).

Combining the information from the two decennial Census files, we can identify about 152 million unique parent–child links for about 81 million children. Note that because of the way that the "Probable" links are identified, it is possible that some of these links represent changes in family structure (marriages, divorces, and creation/dissolution of unmarried partnerships). About 123 million links occur for children with 1 or 2 unique links, while the remaining 29 million occur for children with 3 or more links (these represent cases where parental relationships appear to have changed).

C.2.2 Other Demographic Surveys

The final source of data on parent–child links comes from demographic surveys. These surveys are substantially smaller than the Census, but allow us to identify relationships in non-Decennial year. We use the American Community Survey, which has been conducted monthly since 2001, as our main survey source. The Current Population Survey Annual Social and Economic Supplement (CPS ASEC), which is conducted every March is another potential source of parent–child links. However, the CPS ASEC sample size is substantially smaller than the ACS (200,000 individuals in the CPS-ASEC versus about 5 million in the ACS 1 year files), and thus provides minimal additional information relative to the decennial Census and ACS.

The ACS was conducted as an experimental survey from 2001-2004, with increasing sample size in each year. From 2005–2015, the ACS has consisted of a sample size of about 5 million individuals. The content of the ACS has evolved considerably over this period. In particular, the household relationship question was redesigned for the 2008 ACS. From 2001–2007, the ACS relationship question allowed for 10 categories, with a single "child of householder" category that includes adopted children, biological children and step-children. From 2008–2015, the relationship question was expanded to 13 categories, with separate categories for adopted children, biological children and step-children. As with the Decennial HDF data, the relationship variable in the ACS contains categories for married and unmarried partners of the householder, so we categorize the ACS links as "Certain" (for the householder), "Probable-Married" (for the householder's spouse) and "Probable-Unmarried" (for the householder's unmarried partner). We are able to collect about 22 million parent–child links for about 12 million children.

C.2.3 Combining Relationship Information

The relationship information we have extracted from Decennial Census data and demographic surveys has a substantial degree of overlap. In total, we identify links for over 168 million children.

Note while a vast majority (about 87 percent) of children can be linked to one or two parents, there are a substantial number who are linked to three or more parents.

We initially retain links from each source, to allow for robustness checks on the type of link used (i.e. just using Census links or keeping only "certain" parent–child links). Some source information is included in the data, including the parent and child types from the Census and survey data and the year(s) a link appears in the ACS data. Our main estimating sample uses only the certain links found in the Census and ACS data. In robustness exercises, we additionally include probabilistic links.

C.3 Pollution Exposure at Birth

To analyze the intergenerational effects of pollution exposure, we need to be able to infer the level of ambient air pollution and the changes in EPA policy (designation nonattainment of NAAQS) that parents were exposed to at birth. We do this in three steps. First, we link the set of unique parents identified in the previous section to the Census Numident to obtain date and place of birth. We then obtain monitor-level daily pollution measures from the EPA, which aggregate to the county level, and link these county-level measures to the parents' place of birth. Finally, since the EPA's records of nonattainment designations appears to be incomplete or destroyed, we simulate these nonattainment designations for counties with EPA monitors active in 1969 (before CAAA 1970).

C.3.1 Census Numident Data

Our source of information on the parents' place of birth comes from the Census Numident, which is a derivative product of the SSA Numerical Identification File, and serves as the reference file for the PVS matching algorithm. The Census Numident contains three fields which can be used to infer place of birth, which are transcribed from form SS-5 (application for social security number). The field *pobfin* contains a two digit code for the country of birth for non-native born individuals, and the field *pobst* contains a two character abbreviation for state of birth for all native born US citizens. Both of these fields can be assigned one-to-one with standard geographies (i.e. FIPS codes). The field *pobcity*, however, is slightly more cumbersome. This variable represents the first 12 characters of the place (or county) of birth entered on form SS-5. There is little standardization or cleaning done by SSA or Census for this field, and thus there are numerous misspellings and inconsistencies.

In order to match the information in the *pobcity* with standardized geographies (i.e. county FIPS codes), we take a two-step approach. First, after excluding foreign-born individuals (about 13 million parents), we capitalize on a crosswalk developed jointly by Census researchers and external researchers including Martha Bailey and Reed Walker (described in detail in (Isen et al., 2017)). This crosswalk provides all exact matches (after standardization) and probabilistic matches between *pobcity* entries and unique GNIS place names. A second crosswalk between GNIS places and county FIPS codes allows us to directly match parents to counties exactly. For the remaining cases, we execute a probabilistic matching algorithm. This algorithm assigns a

match by calculating the optimal string alignment (OSA) distance between a *pobcity* entry and a reference list of all county and Census place names, selecting the smallest distance (maximum of 5) within *pobst*. This is essentially the same algorithm as in Voorheis (2017). All told, about 74 percent of native-born parents can be assigned a place of birth using the GNIS crosswalks, and another 23 percent can be matched using our probabilistic matching algorithm, so that about 97 percent of native born parents can be assigned a county of birth.

C.3.2 EPA Monitor Data

With information about the place of parents' birth in hand, we infer the level of pollution exposure experienced by these individuals if we have some information based on the average exposure within their county of birth. To gather this pollution exposure information, we rely on monitor data from the EPA. The EPA has made monitor-level air quality data available via the AQDM API. Our pollutant of interest is particulate matter. For the relevant period of time (around 1970), the primary regulated pollutant was total suspended particles (TSP), defined as the density of particulates less than 100 microns, measured in units $\mu g/m^3$.[19] We thus retrieve all TSP (EPA pollutant code 11101) monitor observations between 1969–1980.

The TSP standard was set based on a 24-hour sampling, and hence the monitor-level data is provided on a daily basis. Our baseline approach to aggregating these daily monitor-level observations is as follows. For each county-day, we calculate the average TSP concentration across all active monitors in that day, which we take as the average exposure to TSP in that county on that day. We then calculate county-level moving average exposure to TSP for each unique birthday between 1969 and 1980 for two periods of interest: the nine months before birth (in utero exposure) and the year after birth (infant exposure).

The EPA's monitoring network expanded dramatically following the passage of CAAA 1970, expanding both the number of counties monitored and the density of monitors within consistently monitored counties. This poses two potential challenges to our baseline measurement approach above. First, some counties will only have observations in the "post-treatment" period in our OLS and IV regressions. Second, even for counties which are consistently monitored, the expansion of the monitor network may result in systematic measurement error—average county exposure will be more precisely measured with more monitors and so the pre-treatment observations are more likely to be mismeasured than the post-treatment observations. To address these issues, we also produce county-level moving averages using a constant set of monitors (the monitors that were active in 1969 or earlier).

C.3.3 Nonattainment Designations

Our empirical strategy relies on identifying the intergenerational effects of pollution exposure at birth using plausibly exogenous variation in TSP exposure that resulted from counties being designated as in nonattainment of the ambient air quality standards in the 1970 CAAA by the EPA. Although the EPA makes nonattainment designations publicly available starting in 1991, and researchers have reconstructed nonattainment designations back to 1980, there appear to be no existing records on which counties were initially designated as being in nonattainment in 1972, the first

[19] This definition was later revised to 10 microns (PM10) and 2.5 microns (PM2.5) standards in 1987 and 1997 respectively.

year in which the 1970 CAAA was in effect. This is because the EPA designated entire air quality regions as in nonattainment, not individual counties, when setting standards authorized under CAAA 1970. In practice, however, the way that these nonattainment designations were applied bound at the county level: the EPA considered an air quality district to be in nonattainment if any of the monitors in the region were in nonattainment. Thus in practice counties with monitors in attainment were not subject to the same regulations as counties that were in nonattainment. Since the TSP air quality standards are known, and as noted in the previous section, we have monitor-level data on the actual level of exposure in the years before the 1970 CAAA was in effect, it is possible to reconstruct which counties would have been in nonattainment.

Nonattainment of the primary air quality standard for TSP set in 1970 CAAA occurs if either a) the annual average (geometric mean) TSP concentration is above 75 $\mu g/m^3$, or b) the second highest daily TSP concentration is above 260 $\mu g/m^3$. We use the monitor-level observations from the previous section to calculate the geometric mean and second highest daily TSP concentration for all counties with at least on monitor in 1970. This allows us to categorize 258 counties as "nonattainment" counties, and 319 counties as "attainment" counties.

C.3.4 Other County Attributes

Estimating the effects of pollution exposure at birth on adult outcomes for parents and intergenerational effects for their children may be confounded by other characteristics of the parents' place of birth, such as weather or economic activity. To this end, we obtain pre-determined (i.e. before the clean air act of 1970) information on county level economic activity from the BEA, and county-level weather information from the National Oceanic and Atmospheric Administration (NOAA).

Following Isen et al. (2017), we obtain information on the economy and population of U.S. counties in 1969 from the Bureau of Economic Analysis' Regional Economic Accounts (1969 is the earliest year for which the BEA publishes regional accounts data). We extract four variables of interest from the regional accounts: total population, total employment, total personal income and total personal transfer income. From these we can construct income per capita and employment-to-population ratio measures; these measures allow us to control for important county-level economic characteristics that may confound the nonattainment-pollution relationship.

Additionally, we obtain information on county-level weather patterns. Temperature and precipitation, in particular, play important factors in the formation of particulate matter emissions and in the suspension of particulate matter in the atmosphere after emission. Importantly, there is evidence both that very low temperatures can increase PM concentrations by emissions (at cold temperatures, internal combustion engines burn fuel less efficiently), while very high temperatures can increase PM concentrations through suspension and atmospheric particle formation (sulfate and nitrate particles form more readily at hot temperatures. Additionally, precipitation decreases PM concentrations by decreasing suspension. Thus we obtain weather-station level data on daily high temperature, low temperature and precipitation from NOAA's Global Historical Climatology Network (GHCN). For each day between 1959 and 1981, we interpolate across the weather station network to each county centroid using inverse distance weighting to obtain a county-day level dataset. We can then calculate the average high/low temperature and number of precipitation days corresponding to the 9 months before birth and the year after birth for each individual.

C.4 ATUS Data

To investigate mechanisms underlying the second generation effect, we will leverage a secondary linked dataset which will allow us to measure both time use for individuals at a point in time, as well as their place of birth and the level of pollution they were exposed to. We do this by linking a subset of respondents to the American Time Use Survey (ATUS) to the Census Numident.

Using the IPUMS public use ATUS data from 2003–2017, we build a series of time use variables which divide the total time spent during the reference day on specific child-enrichment activities (time spent on children's education activities, time spent on children's health activities, time spent reading to a child), as well as broad categories of non-sleep time use (time spent on work, time spent on social activities, time spent on leisure, time spent on education). We then link a subset of the ATUS respondents to the Census Numident to attach place of birth characteristics as follows.

Our linkage strategy relies on the fact that the ATUS sample frame is drawn from the Current Population Survey. Hence it is possible to link ATUS respondents to the CPS on an individual level in the public use data. For the subset of individuals who are in sample and respond to the ASEC, we can link this public use identifier to the internal confidential CPS-ASEC data. The internal CPS-ASEC has had PIKs assigned, so we are then able to link these subset of individuals to the Census Numident by PIK, identifying place of birth and TSP exposure at birth using the same method used for the ACS sample, described above. We further subset this linked sample to individuals born between 1969 and 1980, coinciding with the first generation for the main ACS results. Note that this is a relatively small subsample of ATUS respondents (the final analysis sample for parents aged 25-40 has about 5,000 observations).

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1732	Kabir Dasgupta André Diegmann Tom Kirchmaier Alexander Plum	Heterogeneity in criminal behavior after child birth: the role of ethnicity
1731	Andreas Diemer Tanner Regan	No inventor is an island: social connectedness and the geography of knowledge flows in the US
1730	Hanming Fang Chunmian Ge Hanwei Huang Hongbin Li	Pandemics, global supply chains, and local labor demand: evidence from 100 million posted jobs in China
1729	Ria Ivandić Tom Kirchmaier Ben Linton	Changing patterns of domestic abuse during COVID-19 lockdown
1728	Jonathan Colmer Ralf Martin Mirabelle Muûls Ulrich J. Wagner	Does pricing carbon mitigate climate change? Firm-level evidence from the European Union emissions trading scheme
1727	Tony Beatton Michael P. Kidd Matteo Sandi	School indiscipline and crime
1726	Maximilian v. Ehrlich Henry G. Overman	Place-based policies and spatial disparities across European cities
1725	Gabriel M. Ahlfeldt Thilo N. H. Albers Kristian Behrens	Prime Locations
1724	Benjamin Handel Jonathan Kolstad Thomas Minten Johannes Spinnewijn	The Social Determinants of Choice Quality: Evidence from Health Insurance in the Netherlands

1723	Claudia Hupkau Barbara Petrongolo	Work, Care and Gender During the Covid-19 Crisis
1722	Ross Levine Yona Rubinstein	Selection Into Entrepreneurship and Self-Employment
1721	Sandra McNally	Gender Differences in Tertiary Education: What Explains STEM Participation?
1720	Edoardo di Porto Paolo Naticchioni Vincenzo Scrutinio	Partial Lockdown and the Spread of Covid-19: Lessons From the Italian Case
1719	Swati Dhingra Stephen Machin	The Crisis and Job Guarantees in Urban India
1718	Stephen J. Redding	Trade and Geography
1717	Arun Advani Felix Koenig Lorenzo Pessina Andy Summers	Importing Inequality: Immigration and the Top 1 Percent
1716	Pol Antràs Stephen J. Redding Esteban Rossi-Hansberg	Globalization and Pandemics
1715	Davin Chor Kalina Manova Zhihong Yu	Growing Like China: Firm Performance and Global Production Line Position
1714	Luna Bellani Anselm Hager Stephan E. Maurer	The Long Shadow of Slavery: The Persistence of Slave Owners in Southern Law-Making

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE