The long-term effect of childhood exposure to technology using surrogates*

Sylvia Klosin*		Nicolaj Søndergaard Mühlbach†

DISCLAIMER: This is work in progress and content is likely to change substantially.

This version: May 5 2022

Abstract

We study how childhood exposure to technology at ages 5-15 via the occupation of the parents affects the ability to climb the social ladder in terms of income at ages 45-49 using the Danish micro data from years 1961-2019. Our measure of technology exposure covers the degree to which using computers (hardware and software) is required to perform an occupation, and it is created by merging occupational codes with detailed data from O*NET. The challenge in estimating this effect is that long-term outcome is observed over a different time horizon than our treatment of interest. We therefore adapt the surrogate index methodology, linking the effect of our childhood treatment on intermediate surrogates, such as income and education at ages 25-29, to the effect on adulthood income. We estimate that a one standard error increase in exposure to technology increases the income rank by 2%-points, which is economically and statistically significant and robust to cluster-correlation within families. The derived policy recommendation is to update the educational curriculum to expose children to computers to a higher degree, which may then act as a social leveler.

Keywords: Surrogate index;

JEL Classification:

*The author is grateful to Michael Rosholm and Michael Svarer for helpful comments and suggestions, and to the Department of Economics, Massachusetts Institute of Technology (MIT), the Center for Research in Econometric Analysis of Time Series (CREATES), the Dale T. Mortensen Center, Aarhus University, and the Danish Council for Independent Research (Grant 0166-00020B) for research support.

*Department of Economics, MIT. Email: klosins@mit.edu.

†Department of Economics, MIT, and CREATES. Email: muhlbach@mit.edu.

I. Introduction

Modern technologies have the potential to drastically reshape our economies (Agrawal et al., 2019). The implications for labor markets are, however, not fully understood, nor are the concerns new. As noted by Nobel laureate Wassily Leontief: "*Labor will become less and less important* (...). *More workers will be replaced by machines*" (Leontief, 1952). Throughout history, technological progress has generated cultural anxiety despite being widely considered as the main source of economic growth. The common concern covers technological unemployment stemming from broad substitution of machines for labor (Mokyr et al., 2015). However, transformations have taken place before, and yet, workers continue to have jobs (Autor, 2015). But unlike prior technologies, modern technologies, e.g., artificial intelligence (AI), have the potential to disrupt high-skilled jobs (see, e.g., Darcy et al. (2016) or Dunis et al. (2016)). This raises the question of whether new technologies are fundamentally different from previous technologies in a way that Leontief's prophecy may materialize. Two contrasting beliefs exist (Brynjolfsson and McAfee, 2012, 2014). On the one hand, the pessimistic perspective argues that human labor becomes obsolete as new technology replaces labor tasks. That is, a worker competes with technology. The resulting technological unemployment has been widely studied, e.g., Frey and Osborne (2017) conclude that 47% of U.S. employment is at risk of automation. On the other hand, the optimistic perspective views technology as an enabler that complements certain skills. In this case, a worker is augmented by technology. The transition costs associated with the labor substitution are, in this view, outweighed by the efficiency gains from augmentation. The reason why the two contrasting views co-exist may stem from the lack of high-quality data on the subject (Frank et al., 2019), which is unsatisfactory for policy-makers as the implications are widely different. But Acemoglu and Restrepo (2019b) argue that the debate centers around a false dichotomy and offer an alternative framework for thinking about how technology impacts tasks, productivity, and work through automation. The main idea is that technology creates a displacement effect by replacing workers in tasks that they previously performed, which reduces the demand for labor, wages, and employment. But several countervailing forces could push against the displacement effect, e.g., productivity effects, capital accumulation, deepening of automation, or most importantly, the creation of new tasks leading to a reinstatement effect. Both Acemoglu and Restrepo (2019a) and Acemoglu and Restrepo (2020) build on the idea of two major effects in opposite directions, i.e., the displacement and reinforcement effects, and offer concrete empirical evidence. For instance, Acemoglu and Restrepo (2020) use industrial robots to proxy technology and estimate that one additional robot per thousand workers reduces wages by 0.42% and the employment-to-population ratio by 0.2%-points. More generally, the creation of new tasks does not happen automatically and has not received sufficient focus, evidence suggests (Acemoglu and Restrepo, 2019c). One consequence is rising inequality.

While many studies focus on the impact of technology on employment *within* generations, there is very limited research addressing how these technologies impact the opportunities *between* generations, conceptualized as social mobility. Many studies in sociology and economics focus on social mobility, which captures the association between the social position of an individual's family in her upbringing and the social position of that same individual when she is an adult. The concept of moving between classes is precisely the idea behind *status attainment*, which may be affected by achieved or ascribed factors (Blau and Duncan, 1967; Boudon, 1974). The economic approach to social mobility started by modeling social interactions within families (Becker, 1974; Becker and Tomes, 1976, 1979) and later by measuring the correlation between lifetime earnings of fathers and sons (Solon, 1992; Zimmerman, 1992). More recently, Chetty et al. (2014a) operationalize upward mobility by defining it as the probability that a child reaches the top fifth of the income distribution conditional on having parents in the bottom fifth. Studies on the interplay between inequality and social mobility suggest that inequality lowers mobility: Chetty et al. (2014a) show that social mobility varies substantially across the US; Chetty et al. (2014b) find that children entering the labor market today face the same mobility as children born in the 70s, whereas Chetty et al. (2017) find that the rates of absolute mobility have fallen significantly for children born in the 80s compared to the 40s; Chetty and Hendren (2018a), Chetty and Hendren (2018b), and Chetty et al. (2016) show that the neighborhoods in which children grow up shape their income prospects. Limited research, however, focuses on how technology impacts social mobility.

Specifically, there have been no empirical studies to our knowledge assessing if parent's exposure to technology can act as a social leveler, enabling children from disadvantaged backgrounds to overcome their circumstances. There are various potential mechanisms at work here. When parents are exposed to technology at the job, this is likely to be a topic of conversation over the dinner table and children may inherit and build both interests and knowledge of technology. This could either activate and reinforce certain cognitive abilities to a higher extent or simply evolve to high-demand skills, e.g., programming skills. Related, when parents are exposed to technology, they are more plausibly more likely to privately invest in technologies, e.g., buying computers and other hardware. The children of those parents naturally become more familiar with technology and understand its fundamentals. The interests, knowledge, and understanding push children into technology-related educations and industries, and as wages for STEM majors and in the technology industry tend to be very competitive, this would be a channel via which exposure to technology would increase social mobility. Additionally, if internet access is almost universal, technology democratizes access to knowledge by distributing it more evenly, which could shrink the gap in cultural capital between social classes, and thus may lead to improved social mobility. Cultural capital covers the accumulation of knowledge, behaviors, and skills that can be used to reinforce class differences (Bourdieu, 1973, 1977, 1984).

Of the biggest barriers to study the effect of childhood exposure to technology on social mobility is the lack of high-quality data on the subject (Frank et al., 2019) and the lack of sufficiently long panels of technology exposure and parent-child pairs. In essence, the childhood exposure to technology and the adulthood income are observed over different time horizons and truly far in time from each other. Another data challenge concerns the ability to identify causal effects, which is hard due to unobserved family heterogeneity. It is, however, not until we overcome the data barriers, we can truly address the consequences of technology between generations, and understand the policy implications.

The contribution of this paper is to estimate the impact of technology on social mobility, overcoming the aforementioned data challenges. Getting high-quality measures of parent exposure to technology, we rely on the Danish micro data containing occupational codes at the individual level per year dating back to the early 80s, which we merge with detailed data from the Occupational Information Network (O*NET) that provides measures of hundreds of occupational features.

Our measure of occupational exposure to technology is the degree of working with computers, that is, the degree to which using computers and computer systems to program, write software, process information, etc. is required or needed to perform the occupation. Using this measure, the occupations most exposed to technology are computer systems engineers, computer research scientists, computer programmers, and akin.

Dealing with the lack of sufficiently long panels of parent-child pairs, we adopt and extend the multiple surrogates index framework by Athey et al. (2016) to allow for continuous treatment exposures. That is, we show how to estimate the average derivative of the continuous treatment on the long-term outcome via intermediate variables called (statistical) surrogates. In our study, we observe the exposure to technology at ages 5-15 and estimate the effects on income at ages 45-49 via income, education, job positions, and industries employed at ages 25-29.

To identify any causal effects, we adopt the design in Chetty and Hendren (2018a) by exploiting variation in the age of children when their parents encounter shifting technology exposures to estimate the effects of childhood exposure to technology. The outcomes of children whose family experiences increased exposure may change in proportion to the amount of time they spend growing up with that exposure. Confounding is concerning if e.g., younger kids benefit from more family resources over time. Moreover, because we rely on job shifts over time, we need to control for other changes that happen simultaneously with changes in exposure to technology when one parent shifts job. We achieve this by exploiting the richness of the Danish micro data, where we obtain data on earnings, health proxies, and labor-market experience together with demographic background variables. Additionally, we parameterize occupations by using information from more than 240 occupational attributes from O*NET.

Our findings suggest that childhood exposure to technology significantly impacts the opportunities in adulthood. Specifically, we estimate that a one standard error increase in exposure to technology through the parents' occupations at ages 5-15 leads to an expected rise of more than 2%-points in the income rank at ages 45-49. This has the potential to guide policy-makers; with technology exposure being an enabler, the policy-makers may incentivize children from disadvantaged backgrounds to acquire technology skills via interventions in the education system, thereby improving opportunities for upward mobility. Because the technology exposure captures the use of computer and computer programs to write software, the policy recommendation is to update the educational curriculum to include more computer (software and hardware) classes.

This paper contributes to two strands of literature. First, our results on the effect of childhood exposure to technology on the social mobility suggest the early exposure to technology may act as a social leveler, enabling children from disadvantaged backgrounds to overcome their circumstances. Second, we extend the methodology of multiple surrogates index to allow for continuous treatments and showcase an application.

The rest of this paper is structured as follows. In Section II, we present our extension to the multiple surrogates index framework and provide formal guarantees. Section III introduces the data, and Section IV presents the empirical results, and we discuss the implications of the findings. Section V concludes.

II. Methodology

A. Setup

We consider two non-overlapping samples denoted the Experimental (E) sample and the Observational (O) sample, respectively. We index units from the samples by $i \in [N_E]$ and $i \in [N_O]$, respectively. For convenience, we view the data as one single sample of size $N = N_E + N_O$ equipped with a binary indicator $P_i \in \{E, O\}$, governing to which sample each unit i belongs. Each unit belongs uniquely to a family $j \in [J]$ and N_j is the set of i-indices that belong to jth family with $|N_j| > 1$. When needed, we double-index individuals by both i and j although any particular individual, say i', cannot belong to two families simultaneously.

For experimental units, we observe a single continuous treatment of interest, $W_i^a \in \mathbb{R}$, for ages $a \in [a_1, a_2]$, whereas for observational units, we observe a single long-term outcome, $Y_i^a \in \mathbb{R}$, for ages $a \in [c_1, c_2]$. Thus, the treatment is not observed in the observational sample and the long-term outcome is not observed in the experimental sample. For all units, we observe intermediate outcomes called surrogates, $S_i^a \in \mathbb{R}^K$, for ages $a \in [b_1, b_2]$ and pre-treatment covariates $X_i \in \mathbb{R}^p$ known not to be affected by the treatment. Note that $a_1 < a_2 < b_1 < b_2 < c_1 < c_2$ and that $a_2 - a_1$ could be different from $b_2 - b_1$ and likewise $b_2 - b_1$ could be different from $c_2 - c_1$. In words, we do not necessarily observe the treatment, surrogates, and outcome for the same number of ages (and in particular not for the exact same ages), and the timing of events is that treatment happens strictly before surrogates which occur strictly before the long-term outcomes. Rather than modeling ages separately, we consider averages over ages. Let superscript $[a, b]$ denote the average from age a to b, including both a and b, such that, e.g., $Z_i^{[a,b]} = \frac{1}{b-a+1} \sum_{t=a}^{b} Z_i^t$. For notational convenience, we henceforth drop the subscript and simply use $W_i := W_i^{[a_1,a_2]}$, $S_i := S_i^{[b_1,b_2]}$, and $Y_i := Y_i^{[c_1,c_2]}$.

Following the potential outcomes framework or Rubin Causal Model (Rubin, 1974; Holland, 1986), we posit the existence of a set of potential outcomes $Y_i(w)$ and surrogates $S_i(w)$ for $w \in \mathcal{W} \subseteq \mathbb{R}$. The $Y_i(w)$ is referred to as the individual-level dose–response function, whereas $\mu(w) = \mathbb{E}[Y_i(w)]$ represents the average dose–response function. The object of interest is the average derivative of the dose-response function, namely

$$\tau_0 = \mathbb{E}\left[\frac{\partial Y_i}{\partial W_i} | P_i = E\right], \tag{1}$$

Toy model illustrating family fixed effects For illustrative purposes, imagine for a second that we had access to Y_i and W_i in the same sample. In this case, the simplest model to consider would be

$$Y_{i,j} = \delta_j + \tau W_{i,j} + \gamma X_{i,j} + \varepsilon_{i,j} \tag{2}$$

where δ_j is the family fixed effect, τ measures the effect of a one-unit increase in the average treatment over the ages a_1 to a_2 on the average long-term outcome over the ages c_1 to c_2, and $\varepsilon_{i,j}$ is the error term. Ignoring the family fixed effects, we would estimate (2) by least squares and let δ_j be absorbed into the new error term, $\epsilon_{i,j} = \delta_j + \varepsilon_{i,j}$. Dealing with family fixed effects, however, we rely on differences within families to cancel out δ_j. Let $Z_{\bar{i},j} := \frac{1}{|N_j|} \sum_{i \in N_j} Z_{i,j}$ denote the family average of a given random variable $Z_{i,j}$ and also denote the family-demeaned variable by $\tilde{Z}_{i,j} = Z_{i,j} - Z_{\bar{i},j}$. The family-demeaned version of Eq. (2) would be

$$\tilde{Y}_{i,j} = \tau \tilde{W}_{i,j} + \gamma \tilde{X}_{i,j} + \tilde{\varepsilon}_{i,j} \tag{3}$$

We do not, however, observe the treatment and the long-term outcome in the same sample, for which reason we introduce the multiple surrogates index (MSI).

B. The multiple surrogates index

For units in the experimental sample, let $f_{W_i|X_i}(w|x)$ for all $w \in \mathcal{W}$ denote the assignment mechanism defined as the conditional probability density function of each treatment level given the pre-treatment covariates $X_i = x$. To introduce the MSI, we present our assumptions needed for our estimator to recover our object of interest.

Assumption 1 (Overlap). *For all $x \in \mathcal{X}$, the conditional probability density function of receiving any treatment $w \in \mathcal{W}$ is bounded away from zero:*

$$f_{W_i|X_i}(w|x) > 0 \quad \forall w \in \mathcal{W}, x \in \mathcal{X} \tag{4}$$

Assumption 2 (Unconfoundedness). *The assignment mechanism is strongly unconfounded such that for each unit $i \in [N]$:*

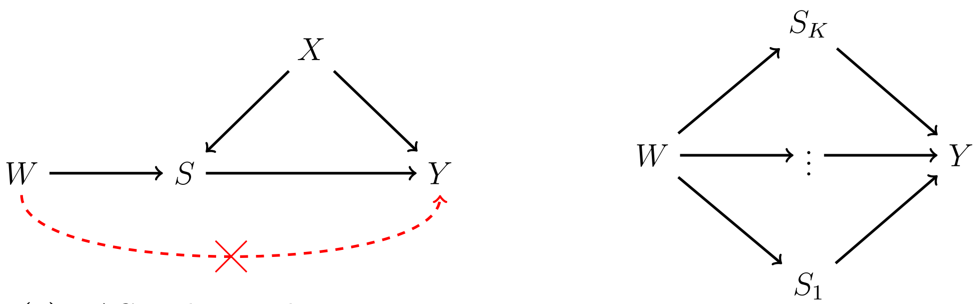
$$W_i \perp\!\!\!\perp (Y_i(w), S_i(w)) | X_i, P_i = E \quad \forall w \in \mathcal{W} \tag{5}$$

Assumptions 1–2 are standard in the causal inference literature and implies that if Y_i was observed for the experimental units, we could estimate the average causal derivative of the treatment W_i on the outcome Y_i by adjusting for pre-treatment covariates, X_i.

Assumption 3 (Surrogacy). *The treatment, W_i, and long-term outcome, Y_i, are conditionally independent:*

$$W_i \perp\!\!\!\perp Y_i | S_i, X_i, P_i = E \tag{6}$$

Assumption 3 defines the key property between W_i, S_i, and Y_i because it assumes that the causal link between the treatment and long-term outcome is fully captured by the surrogates. In Figure 1, we illustrate the implications of Assumption 3. To the left, Figure 1a illustrates the assumption using a single surrogate, where W cannot directly affect Y only indirectly through S. However, X is allowed to affect both S and Y, but cannot be affected by W. To the right, Figure 1b makes Assumption 3 more plausible by including multiple surrogates. Note that the role of X is ignored in Figure 1b for illustratory purposes.



(a) DAG with a single surrogate

(b) DAG with multiple surrogates

Figure 1: Illustrating the multiple surrogates index method

Notes: This figure shows

The last of the fundamental assumptions asserts the comparability of the samples.

Assumption 4 (Comparability). *The conditional distribution of the outcome Y_i is the same in both samples:*

$$Y_i | S_i, X_i, P_i = O \sim Y_i | S_i, X_i, P_i = E \tag{7}$$

The simplest estimator A natural empirical analog of (1) is

$$\hat{\tau} = \mathbb{E}_N \left[\frac{\partial Y_i}{\partial W_i} | P_i = E \right], \tag{8}$$

which we cannot directly compute because Y_i is not observed in the experimental sample. However, one may rewrite Eq. (8) as

$$\begin{aligned}
&\mathbb{E}_N \left[\frac{\partial Y_i}{\partial W_i} | P_i = E \right] \\
&= \mathbb{E}_N \left[\frac{\partial Y_i}{\partial S_i} \times \frac{\partial S_i}{\partial W_i} | P_i = E \right] \\
&= \mathbb{E}_N \left[\frac{\partial Y_i}{\partial S_i} | P_i = E \right] \times \mathbb{E}_N \left[\frac{\partial S_i}{\partial W_i} | P_i = E \right] + \mathbb{COV}_N \left(\frac{\partial Y_i}{\partial S_i}, \frac{\partial S_i}{\partial W_i} | P_i = E \right) \\
&= \mathbb{E}_N \left[\frac{\partial Y_i}{\partial S_i} | P_i = O \right] \times \mathbb{E}_N \left[\frac{\partial S_i}{\partial W_i} | P_i = E \right] + \mathbb{COV}_N \left(\frac{\partial Y_i}{\partial S_i}, \frac{\partial S_i}{\partial W_i} | P_i = E \right) \qquad (9)
\end{aligned}$$

The first equal sign is due to Assumption 3, where we estimate $\frac{\partial Y_i}{\partial S_i}$ separately from $\frac{\partial S_i}{\partial W_i}$. The second equal sign uses the definition of covariance. Assumption 4 gives the the third equal sign because the sample E and O are comparable. The covariance term is not estimable because Y_i and W_i are not observed in the same sample. But assuming that either Y_i is homogeneous in S_i or that S_i is homogeneous in W_i makes $\mathbb{COV}\left(\frac{\partial Y_i}{\partial S_i}, \frac{\partial S_i}{\partial W_i}|P_i = E\right) = 0$. Under this assumption, a simple estimator reads

$$\hat{\tau} = \mathbb{E}_N\left[\frac{\partial Y_i}{\partial S_i}|P_i = O\right] \times \mathbb{E}_N\left[\frac{\partial S_i}{\partial W_i}|P_i = E\right]. \tag{10}$$

Last, under Assumption 1–2, the estimator in (10) identifies the causal average derivative.

Toy model illustrating the treatment effect We illustrate how to operationalize (10) under linear models. For the experimental sample, we consider a linear model for each $k \in [K]$ surrogate

$$\tilde{S}_{i,k} = \theta_k \tilde{W}_i + \gamma \tilde{X}_i + \tilde{v}_{i,k}, \tag{11}$$

where we ignore the family subscript j for notational convenience. For the observational sample, we consider a single linear model

$$Y_i = \beta_0 + \sum_{k=1}^{K} \beta_k S_i + \delta X_i + v_i. \tag{12}$$

Combining the parameters from the experimental stage (θ_k for $k \in [K]$) with the parameters from the observational stage (β_k for $k \in [K]$), the treatment effect analogously to Eq. (1) is given by

$$\tau = \sum_{k=1}^{K} \theta_k \times \beta_k \tag{13}$$

Estimating the models in Eq. (11) and Eq. (12) by least squares and saving the estimates $\hat{\theta}_k, \hat{\beta}_k$ for $k \in [K]$, the empirical analog to (13) is $\hat{\tau} = \sum_{k=1}^{K} \hat{\theta}_k \times \hat{\beta}_k$.

Toy model illustrating the covariance assumption We illustrate when the covariance between the two derivatives is zero by considering two simple linear models in the experimental and observational stages in Eq. (14) and (15), respectively,

$$S_i = \theta_1 W_i + \theta_2 W_i^2 + \gamma X_i + \psi X_i W_i + u_i \tag{14}$$

$$Y_i = \beta_1 S_i + \beta_2 S_i^2 + \delta X_i + \lambda X_i S_i + v_i, \tag{15}$$

where we have included higher-order terms of W_i and S_i, respectively. Then, Eq. (9) reads

$$\begin{aligned}
&\mathbb{E}_N\left[\frac{\partial Y_i}{\partial S_i}|P_i = O\right] \times \mathbb{E}_N\left[\frac{\partial S_i}{\partial W_i}|P_i = E\right] + \mathbb{COV}_N\left(\frac{\partial Y_i}{\partial S_i}, \frac{\partial S_i}{\partial W_i}|P_i = E\right) \\
&= \mathbb{E}_N\left[\beta_1 + 2\beta_2 S_i + \lambda X_i|P_i = O\right] \times \mathbb{E}_N\left[\theta_1 + 2\theta_2 W_i + \psi X_i|P_i = E\right] \\
&+ \mathbb{COV}_N\left[2\beta_2 S_i + \lambda X_i, 2\theta_2 W_i + \psi X_i|P_i = E\right]
\end{aligned} \tag{16}$$

The covariance term in Eq. (16) equals zero if and only if either $\beta_2 = \lambda = 0$ or $\theta_2 = \psi = 0$, meaning that either W_i has no higher-order effect on S_i or S_i has no higher-order effect on Y_i. In practice, this assumption is highly plausible because it does allow one the models to be heterogeneous in its inputs. We would often assume that Y_i depends non-linearly on S_i, which makes us assume that S_i depends linearly on W_i.

Our preferred estimator However, we can incorporate the estimation of the covariance term by considering an alternative estimator, which acts as our preferred choice. To introduce our preferred estimator, we define the following conditional expectations

$$\gamma_1(s, x) = \mathbb{E}\left[Y_i|S_i = s, X_i = x, P_i = O\right], \tag{17}$$

$$\gamma_2(w, x) = \mathbb{E}\left[\gamma_1|W_i = w, X_i = x, P_i = E\right]. \tag{18}$$

We will argue that $\tau_0 = \mathbb{E}\left[\frac{\partial \gamma_2(W_i, X_i)}{\partial W_i}\right]$, which leads to our estimator of τ_0 being

$$\hat{\tau} = \mathbb{E}\left[\frac{\partial \hat{\gamma}_2(W_i, X_i)}{\partial W_i}|P_i = E\right]. \tag{19}$$

Estimating Eq. (19), we follow the steps below.

1. Estimate Eq. (17) using the observational sample. Specifically, regress Y_i on $b(S_i, X_i)$ and denote the fitted regression by $\hat{\gamma}_1(w, x)$.

2. Generate predictions from $\hat{\gamma}_1(w, x)$ using the experimental data by plugging in observed values of W_i and X_i into the estimated model.

3. Estimate Eq. (18) using the experimental sample. Specifically, regress the predictions $\hat{\gamma}_1(W_i, X_i)$ on W_i and X_i.

4. Obtain $\hat{\tau}$ by differentiating $\hat{\gamma}_2(W_i, X_i)$ with respect to W_i using the experimental sample.

C. Identification

Given Assumptions 1–4, we show that regression Y_i on W_i is equivalent to regressing $\mathbb{E}[Y_i|S_i = s]$ on W_i, where we ignore X_i to economize on notation. That is, we show that

$$\arg\min_{\tau_1 \in \mathbb{R}} \left\{ \mathbb{E}\left[(Y_i - \tau_1 W_i)^2 \right] \right\} = \arg\min_{\tau_2 \in \mathbb{R}} \left\{ \mathbb{E}\left[(\mathbb{E}[Y_i|S_i] - \tau_2 W_i)^2 \right] \right\}. \tag{20}$$

Using the least squares solution, Eq. (20) rewrites to

$$\frac{\mathbb{COV}(Y_i, W_i)}{\mathbb{V}(W_i)} = \frac{\mathbb{COV}(\mathbb{E}[Y_i|S_i], W_i)}{\mathbb{V}(W_i)}. \tag{21}$$

To show that Eq. (21) holds, we rewrite its LHS to

$$\begin{aligned} \frac{\mathbb{COV}(Y_i, W_i)}{\mathbb{V}(W_i)} &= \frac{\mathbb{E}[\mathbb{COV}(Y_i, W_i|S_i)] + \mathbb{COV}(\mathbb{E}[Y_i, S_i], \mathbb{E}[W_i, S_i])}{\mathbb{V}(W_i)} \\ &= \frac{\mathbb{COV}(\mathbb{E}[Y_i, S_i], \mathbb{E}[W_i, S_i])}{\mathbb{V}(W_i)}, \end{aligned} \tag{22}$$

where the first equal sign is due to the law of total covariance and the second to Assumption 3. The final step is to rewrite $\mathbb{E}[W_i, S_i]$ to W_i in the covariance term. Let $W_i = \mathbb{E}[W_i|S_i] + u_i$. First, $\mathbb{COV}(\mathbb{E}[W_i, S_i], u_i) = 0$ because no part of u_i can be correlated with S_i. Then, also by Assumption 3, $\mathbb{COV}(\mathbb{E}[Y_i, S_i], u_i) = 0$, which gives Eq. (21) and then Eq. (20).

Standard errors We construct standard errors of Eq. (19) using Generalized Method of Moments (GMM). In the following, we let $D = (Y, S, W, X)$ denote the quadruple of random variables considered. We consider a system of two equations with two parameters, $\theta = (\theta_1, \theta_2)'$, as in

$$\begin{aligned} Y_i &= S_i'\theta_1 + v_i \\ S_i'\theta_1 &= W_i'\theta_2 + u_i. \end{aligned} \tag{23}$$

Based on Eq. (23), the moment conditions read

$$\begin{aligned} g(D_i, \theta) &= \begin{bmatrix} g_1(D_i, \theta) \\ g_2(D_i, \theta) \end{bmatrix} \\ &= \begin{bmatrix} S_i'(Y_i - S_i'\theta_1) \\ W_i'(S_i'\theta_1 - W_i'\theta_2) \end{bmatrix}, \end{aligned} \tag{24}$$

with orthogonality condition $\mathbb{E}[g(D_i, \theta)] = 0$. The Jacobian then follows as

$$\begin{aligned} G &= \mathbb{E}[\nabla_\theta \{g(D_i, \theta)\}] \\ &= \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial g_1(D_i,\theta)}{\partial \theta_1} & \frac{\partial g_1(D_i,\theta)}{\partial \theta_2} \\ \frac{\partial g_1(D_i,\theta)}{\partial \theta_2} & \frac{\partial g_2(D_i,\theta)}{\partial \theta_2} \end{bmatrix} \\ &= \begin{bmatrix} -S_i'S_i & 0 \\ W_i'S_i & -W_i'W_i \end{bmatrix}, \end{aligned} \tag{25}$$

The influence function is then given by $\psi(D_i, \theta) = G^{-1}g(D_i, \theta)$, which equals ...

$$\begin{aligned} \psi(\theta) &= \begin{bmatrix} \psi_1(\theta) \\ \psi_2(\theta) \end{bmatrix} \\ &= \begin{bmatrix} G_{11}^{-1} g_1(\theta) \\ G_{22}^{-1}(g_2(\theta) + (-G_{21})\psi_1(\theta)) \end{bmatrix} \end{aligned} \tag{26}$$

Consequently, using influence functions, the asymptotic variance of $\hat{\theta}$ reads

$$\hat{\mathbb{V}}\left(\hat{\theta}\right)=\frac{1}{N}\left(\frac{1}{N}\sum_{i=1}^{N}\psi\left(D_i,\hat{\theta}\right)\psi\left(D_i,\hat{\theta}\right)'\right) \tag{27}$$

D. Simulation evidence

We demonstrate the performance of our preferred estimator through a Monte Carlo simulation exercise. We use the data generating process (DGP) defined below to create 10,000 different samples. In each sample, we calculate $\hat{\tau}$ by using the estimator given in equation (19). The distribution of these 10,000 $\hat{\tau}$ values is given in Figure 2.

Data-generating process We consider a single pre-treatment covariate $X \sim N(0,1)$. Our treatment assignment is not random but a function of X and is distributed normally, $D \sim .3X + N(0,1)$. We consider two surrogate variables S_1 and S_2. Both are normally distributed $S_1 \sim 2D + XD + N(0,1)$ and $S_2 \sim 3D + N(0,1)$. Our outcome model is then

$$Y \sim 4S_1 + XS_1 + S_2 + N(0,1) \tag{28}$$

This means that our target τ_0 is equal to

$$\begin{aligned}
\tau_0 &= \mathbb{E}\left[\frac{\partial Y_i}{\partial W_i}|P_i = E\right] \\
&= \mathbb{E}\left[\frac{\partial}{\partial W_i}\left(4S_{i,1} + X_i S_{i,1} + S_{i,2} + \varepsilon_i\right)|P_i = E\right] \\
&= \mathbb{E}\left[\frac{\partial}{\partial W_i}\left(4\left(2D_i + X_i D_i + u_i\right) + X_i\left(2D_i + X_i D_i + u_i\right) + \left(3D_i + v_i\right) + \varepsilon_i\right)|P_i = E\right] \\
&= \mathbb{E}\left[8 + 4X_i + 2X_i + X_i^2 + 3|P_i = E\right] \\
&= \mathbb{E}\left[11 + 6X_i + X_i^2|P_i = E\right] \\
&= 12
\end{aligned} \tag{29}$$

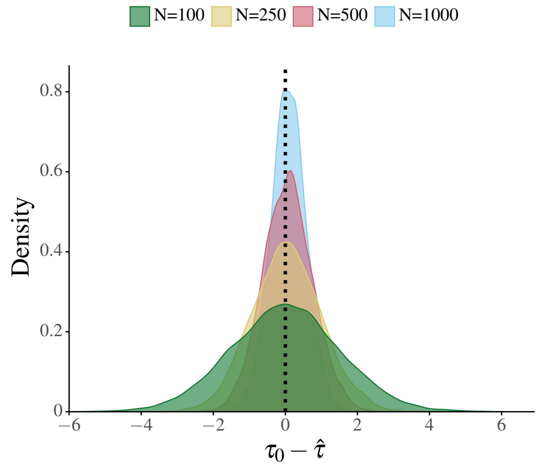


Figure 2: Centered distribution of $\hat{\tau}$

Notes: This figure shows the distribution of $\hat{\tau}$ from a Monte Carlo simulation. This shows the distribution of 10,000 estimates of $\hat{\tau}$. Each $\hat{\tau}$ is estimated using a varying number of observations from $N = 100$ to $N = 1000$.

III. Data

We start by describing the data unique to both the experimental and observational sample, after which we describe the data common to the samples.

A. Experimental sample

The experimental sample is used to estimate the effects of childhood exposure to technology on *expected* long-term income through several intermediate surrogates. Individuals represent cohorts born between 1986 and 1990, leading to five cohorts. We sample the exposure to technology at ages 5-15, which occurs in the years 1991-2005. For instance, the 1986 cohort is sampled between 1991-2001 (when they are 5-15 years of age) to get the treatment exposure.

Treatment The treatment of interest is exposure to technology, which we measure at the occupational level of the parents. By exposure to technology, we mean the occupational degree of *interacting with computers*, which is a work activity defined by "using computers and computer systems (including hardware and software) to program, write software, set up functions, enter data, or process information.". Specifically, for each age of child i and for each parent, we observe the occupation of the parent and attach to it a measure of technology exposure, namely the degree to which interacting with computers is required or needed to perform the occupation.[1] We use the maximum of the parents' technology score per age of the child.

Additional controls In addition to the (maximum) technology score of the parents' occupation in a given year, we observe all of the 244 O*NET attributes at the occupational level as well as the commonly-used tasks measures (abstract, manual, and routine tasks) by Autor et al. (2003, 2006); Acemoglu and Autor (2011); Autor and Dorn (2013). We cannot, however, include all O*NET attributes because the estimation would become too high-dimensional and overfitting would be likely to occur. As an alternative, we include principal components of the O*NET attributes. Specifically, we use the K-means clustering algorithm to create ten clusters of O*NET attributes and locate the one cluster that contains our treatment variable (the work activity capturing the degree of interacting with computers). Then, we apply PCA on all the variables from the remaining nine clusters and extract the first five principal components, which act as a medium-dimensional control for occupational attributes. Further, we characterize the financial situation of the parents by their total personal income, which contains their labor and capital income. We observe the parents' employment status and report the amount of sickness benefits they have received from the government. We follow Gustman and Steinmeier (2018) in including a large set of health variables to deal with the measurement issues of health that are well-established in the literature (see, e.g., Stern (1989); Bound (1991)). Specifically, we characterize the parents' health status by the number of hospitalizations, and the tariffs for visits at the general practitioner, psychiatry, physiotherapy, and surgery, which are summarized as total health expenses.

B. Observational sample

The observational sample is used to study the relationship between the long-term outcome and the intermediate surrogates. Essentially, the observational sample

[1] Technically, we observe the Danish version of the International Standard Classification of Occupations (ISCO) for each occupation, which we first crosswalk to the Standard Occupational Classification (SOC) system and then we crosswalk this to O*NET codes.

Table 1: Data overview

		Treatment exposure		Short-term outcomes		Long-term outcomes	
Sample	Cohorts	Years	Ages	Years	Ages	Years	Ages
O	1961-1970			1986-1999	25-29	2006-2019	45-49
E	1986-1990	1991-2005	5-15	2011-2019	25-29		

Notes: This table provides an overview of the samples used in this study. In particular, the table displays what cohorts, years, and ages that sampled in both the observational and the experimental sample, respectively.

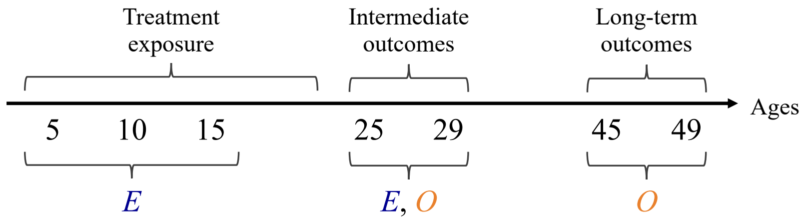


Figure 3: Data overview

Notes: This figure shows the age range included in the observational sample (O) and the experimental sample (E) governing the treatment exposure and the intermediate and long-term outcomes.

allows us to learn the statistical relationship between outcomes at different stages in life. Specifically, individuals come from cohorts born between 1961 to 1970, leading to 10 cohorts. The primary outcome is the (average) income percentile rank at 45-49 years of age, which we sample over the years 2006-2019. Specifically, the 1961 cohort is sampled in 2006-2010 (when they are 45-49 years of age), whereas the 1970 cohort is sampled in 2015-2019.

C. Common to the experimental and observational samples

We observe the surrogates in both samples when the individuals are 25-29 years of age. For the experimental units, this happens in the years 2011-2019. For the observational units, the surrogates are sampled between 1986-1999. We summarize the samples in Table 1 and present a graphical overview in Figure 3.

For each individual, the surrogates include income at each age as well as the average between the ages 25-29, the education level (i.e., lower secondary, upper secondary, short-cycle tertiary, bachelor's degree, or master's degree) and field (i.e., generic, humanities, health, business, engineering, or science) at age 29, the most common job position (i.e., outside the labor force or unemployed, employee, self-employed, manager, or director) and industry (i.e., industries defined by The North American Industry Classification System (NAICS)) between the ages 25-29, and the average sickness benefits between the ages 25-29. In addition, we obtain pre-treatment demographic covariates such as sex, citizenship, and origin. As many algorithms require categorical data to be transformed into numerical data before the data can be used as input, we use one-hot encoding to transform each of the categorical variables.

D. Descriptive statistics

We continue this section by reporting some summary statistics. Specifically, we report the frequency of the categorical variables in Table 2 and the mean of the numerical variables in Table 3. For both tables, we distinguish between the experimental and the observational sample. In Table 3, the surrogates are denoted "Income (age 25)" to "Income (age 29)" as well "Income (age 25-29)" and also "Sickness benefits (age 25-29)". The long-term outcome variable is denoted "Income (age 45-49)". We include a few extra variables in the experimental sample (i.e., "Experience", "Health expenses", "Hospital admissions", "Income", and "Sickness benefits") to control for everything else that potentially happens with job shifts and these variables are measured in the say way as the treatment variable. We do not provide statistics for the principal components as they are standardized by default.

In Table 2, the most notable differences between the two samples are the educational level and field, where the younger cohorts (the experimental sample) are better educated and more often educated (and employed) within health and humanities compared to the older cohorts (the observational sample). For the numerical variables in Table 3, the differences appear negligible. Overall, as we will explore in the next subsection, the differences have little to no statistical impact as only a very limited number of individuals are statistical highly likely (more than 90%) to belong to a particular sample conditional on the surrogates and pre-treatment covariates, for which reason we conclude that there is a reasonable overlap between the samples.

Table 2: Descriptive statistics of categorical variables

Variable	Category	Experimental sample	Observational sample
Cohort	1961		8.6%
	1962		8.9%
	1963		10.1%
	1964		11.0%
	1965		12.0%
	1966		12.4%
	1967		10.9%
	1968		9.2%
	1969		8.4%
	1970		8.5%
	1986	21.4%	
	1987	21.0%	
	1988	20.0%	
	1989	18.7%	
	1990	18.8%	
Sex	Female	48.5%	48.2%
	Male	51.5%	51.8%
Origin	Danish	91.7%	98.5%
	non-Danish	8.3%	1.5%
Citizenship	Danish	97.9%	99.3%
	non-Danish	2.1%	0.7%
Education level	Lower secondary	13.7%	24.6%
	Upper secondary	38.9%	52.3%
	Short-cycle tertiary	5.1%	4.5%
	Bachelor	24.2%	13.1%
	Master	18.1%	5.6%
Education field	Generic	21.5%	34.1%
	Humanities	18.6%	9.1%
	Health	18.0%	8.5%
	Business	17.7%	21.9%
	Engineering	17.7%	22.2%
	Science	6.4%	4.2%
Industry	Agriculture	1.3%	3.2%
	Construction	7.1%	6.5%
	Finance	13.2%	10.3%
	Health	23.0%	15.1%
	Humanities	10.8%	7.7%
	Manufacturing	8.0%	23.2%
	Public	11.7%	10.8%
	Trade	25.0%	23.1%
Job position	Inactive	15.7%	10.2%
	Employee	58.6%	72.6%
	Self-employed	2.1%	3.0%
	Manager	23.0%	14.1%
	Director	0.7%	0.2%

Notes: This table reports the frequencies of the categorical variables for both the experimental and observational sample.

E. Ensuring overlap between samples

Assumption 4 states that the experimental and observational samples are comparable in terms of the conditional distribution of the outcome Y_i. This assumption is untestable because Y_i is not observed in the experimental sample. But as an

Table 3: Descriptive statistics of numerical variables

Variable	Experimental sample	Observational sample
Experience	0.82	
Health expenses	228.45	
Hospital admissions	0.15	
Income	66600.15	
Income (age 25)	31076.07	36040.45
Income (age 26)	34277.33	38193.09
Income (age 27)	38073.63	40210.05
Income (age 28)	41753.91	42146.73
Income (age 29)	45254.06	43969.96
Income (age 25-29)	38100.52	40126.49
Income (age 45-49)		421538.70
Sickness benefits	1002.57	
Sickness benefits (age 25-29)	793.90	874.73
Treatment	0.20	

Notes: This table reports the mean of the numerical variables for both the experimental and observational sample.

robustness assessment and sample selection strategy, we restrict our attention to individuals that statistically *could* belong to either sample in terms of propensity scores.[2] Specifically, define the probability of belonging to the observational sample by $\rho_i = \Pr(P_i = O | S_i, X_i)$. We estimate ρ_i via a (regularized) logistic regression by regressing the indicator $\mathbb{1}\{P_i = O\}$ on $b(S_i, X_i)$, where $b(\cdot)$ is a high-dimensional dictionary and we use the fitted values $\hat{\rho}_i$ as an estimate of the propensity score. The distribution of the estimated propensity scores follow from Figure 4. We keep individuals in the experimental sample if $\hat{\rho}_i > \xi$ and individuals in the observational samples if $\hat{\rho}_i < 1 - \xi$, where ξ acts as a propensity score threshold and is fixed at 0.1. As a consequence of extreme propensity scores, we remove 1697 (0.84%) and 2017 individuals (0.48%) from the experimental and observational sample, respectively. Therefore, we emphasize that this sample selection strategy has almost no effect on the sample as a very limited number of individuals are removed due to extreme propensity scores. This ultimately supports Assumption 4.

[2] Traditionally, the propensity score is the probability of belonging to the treated sample as opposed to the control sample in binary treatment studies. We slightly misuse the terminology and use propensity scores as the probability of belonging to the observational sample as opposed to the experimental sample.

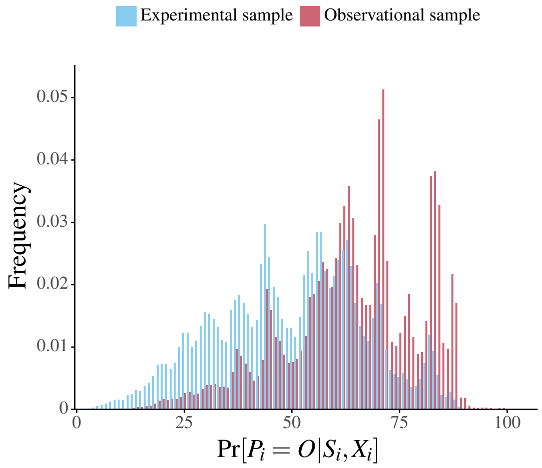


Figure 4: Distribution of propensity score estimates

Notes: This figure shows the distribution of propensity scores estimated by running a regularized logistic regression of $\mathbb{1}\{P_i = O\}$ on $b(S_i, X_i)$, where $b(\cdot)$ is a dictionary.

IV. Empirical Analysis

A. The benefits of multiple surrogates

The idea of using a single intermediate variable as a surrogate to estimate long-term effects is not new and dates back to Prentice (1989) and has been studied by, e.g., Begg and Leung (2000); Frangakis and Rubin (2002). We will, however, argue that the use of multiple surrogates is strongly preferred whenever multiple surrogates are available to the researcher, which is also emphasized by Athey et al. (2016). First, the interpretation of the marginal effect of any single surrogate depends on whether the effect is unconditional or conditional. In our application, we use income at ages 25-29 and the educational attainment at age 29 as surrogates, but income at younger ages plays a very different role in explaining income at older ages when income is conditional on education and when it is not. Ceteris paribus, higher income at younger ages *should* lead to higher income at older ages but highly-educated individuals may have low earnings in their late 20s although they tend to receive high earnings in their late 40s. Second, the use of multiple surrogates rather than one makes Assumption 3 much more likely to hold empirically because the treatment of interest has more channels to (indirectly) affect the long-term outcome variable. Third, the usefulness of the surrogates

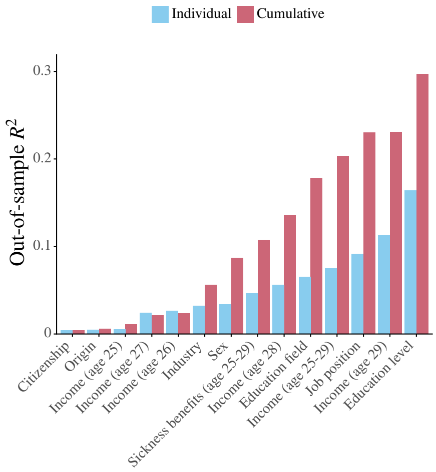


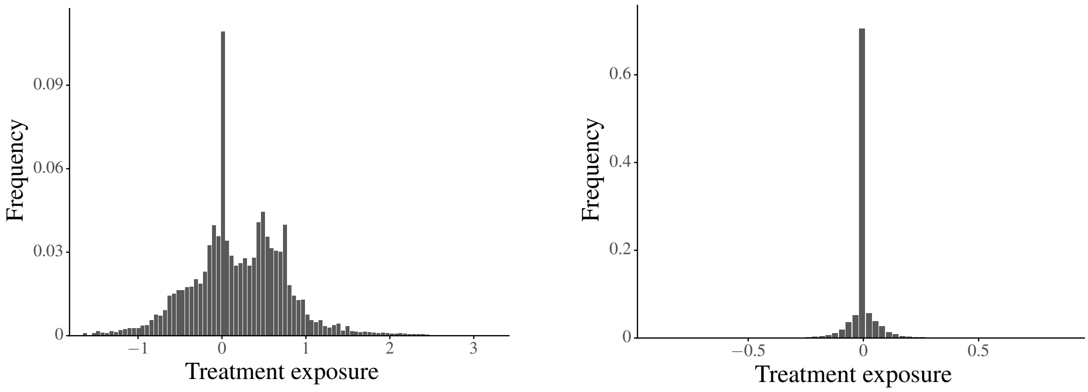
Figure 5: Out-of-sample R^2 for surrogates separately and combined

Notes: This figure shows the out-of-sample R^2 for surrogates separately and combined evaluated on 120,851 test individuals not used to estimate the model parameters. The surrogates are ranked by the individual predictive ability. The last column to the right represents the estimation using all surrogates jointly.

index framework depends strongly on the ability of the surrogates to precisely predict the long-term outcome. Mathematically, the lower the generalization error, $\mathbb{E}\left[\mathcal{L}\left(Y, \gamma_1\left(S, X\right)\right)\right]$, the more useful the framework is, where $\mathcal{L}\left(\cdot, \cdot\right)$ is some loss function and the expectation is taken over an independent test sample. Fortunately, we can assess the predictability of the long-term outcome empirically by each surrogate and jointly. Specifically, we split the observational sample into a training and a test sample using a 75%/25% split, leading to 362,550 and 120,851 individuals in each sample, respectively. Then, for each surrogate $j \in [K]$, we estimate $\gamma_{1,j}\left(s, x\right) = \mathbb{E}\left[Y_i | S_{i,k} = s, X_i = x, P_i = O\right]$ using least squares and compute the out-of-sample R^2 using the test sample as

$$1 - \frac{\sum_{i \in \mathcal{D}_{\text{test}}} \left(Y_i - \hat{\gamma}_{1,j}\left(S_{i,k}, X_i\right)\right)^2}{\sum_{i \in \mathcal{D}_{\text{test}}} \left(Y_i - \overline{Y}_{\mathcal{D}_{\text{train}}}\right)^2}, \tag{30}$$

where $\mathcal{D}_{\text{train}}$ and $\mathcal{D}_{\text{test}}$ are sets containing the indices of the training and test sample, respectively, and $\bar{Y}_{\mathcal{D}_{\text{train}}} = \left|\mathcal{D}_{\text{train}}\right|^{-1} \sum_{i \in \mathcal{D}_{\text{train}}} Y_i$ is the average of the long-term outcome using the training indices. In Figure 5, we plot the individual and cumulative out-of-sample R^2 together with the out-of-sample R^2 stemming from using all the surrogates jointly.



(a) Raw treatment exposure

(b) Family-demeaned treatment exposure

Figure 6: Distribution of the treatment exposure

Notes: This figure shows the treatment of the childhood exposure to technology averaged of the aged 5 to 15. The LHS (Figure 6a) shows the raw treatment exposure, where the RHS (Figure 6b) shows the family-demeaned analog, which is the one used in all regressions.

Figure 5 suggests that using all the available surrogates strongly outperforms any single (or cumulative) surrogate in terms of predicting the long-term outcome. In fact, the highest out-of-sample R^2 from any single surrogate, which is obtained using the education level, only reaches 16.4% compared to 29.7% using all surrogates jointly. As also pointed out by Athey et al. (2016), this ultimately leads to much lower standard errors when estimating the average effect of the treatment on the long-run outcome.

B. Assessing the treatment

Identification of the effect of exposure to technology ultimately depends on the variation in parental occupation and job shifts over time as we remove family fixed effects. To assess the variation in the treatment exposure, we show two histograms in Figure 6, where Figure 6a shows the distribution of the raw treatment exposure and Figure 6b shows the distribution of the family-demeaned treatment exposure, that is, the actual treatment exposure used in the estimation.

As seen from Figure 6b, roughly 70% of the individuals are only mildly if at all exposed to variation in the treatment exposure within the family because their parents barely change jobs over the years in which we measure the treatment (that is, over the ages from 5 to 15 of the child). This means that identification comes from the remaining 30% of the sample, which corresponds to approximately 55,700 individuals. We do, however, believe that this high number of individuals is still sufficient to estimate the effects accurately.

C. Intermediate effects

Although our preferred estimator does not implicitly model the relationship between the treatment exposure and each of the surrogates, we may still assess the intermediate effect of the treatment on the surrogates following the procedure outlined for the simplest estimator, which turns out to also advocate for the use of multiple surrogates. Essentially, we regress each (standardized) surrogate on the treatment, pretreatment variables, and the other occupational control variables and show the coefficient on the treatment variable in Figure 7 with confidence intervals. This approach suggests that childhood exposure to technology at ages 5-15 has a positive impact on the income, educational attainment, and job position in the late 20s. On the contrary, childhood exposure to technology tend to discourage working in the humanities industry and improve health conditions by reducing the amount of sickness benefits received in the late 20s. In terms of statistical significance, these effects are, however, not strong individually and only a few of the effects are significant at conventional levels. This further urges the use of multiple surrogates.

D. Long-term effects

Average effects Using our preferred estimator in Eq. (19), we will next estimate the long-term effects of childhood exposure to technology on income in the late 40s via surrogates from the late 20s. We consider five different specifications that involve different control variables. The baseline specification has no control variables and Eq. (18) is therefore simply a regression of the predicted long-term income rank on childhood exposure to technology. The second specification adds a few demographic variables such as sex, citizenship, and origin as controls. The third specification adds further labor-market experience and three proxies for health as controls. To control for all other changes that happen with job shifts, the fourth specification adds parental income and the first three principal components of the O*NET occupation attributes. Finally, the fifth specification considers additionally two principal components of the O*NET occupation attributes,

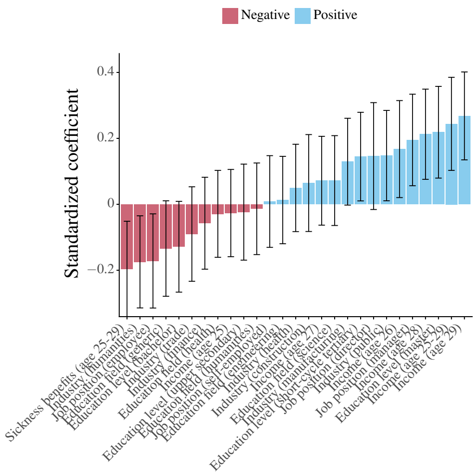


Figure 7: Effects of the treatment on each surrogate

Notes: This figure shows the average effect (as measured by the standardized regression coefficients) of childhood exposure to technology on each of the surrogates. The surrogates are ranked by the effects from most negative to most positive. Negative effects are colored by red, whereas positive effects are colored by blue. Confidence intervals are shown by the black error bars.

that is, the fourth and fifth. Adding more principal components to control for occupational shifts do not significantly alter the results.[3] The regression results from all specifications follow from Table 4. All standard errors are robust to cluster correlation at the family level.

The three most simple specifications (columns 1-3 in Table 4) suggest an average effect of increasing childhood exposure to technology on the income rank in the late 40s of slightly more than 1%-point, which is statistically significant at 10% using cluster-robust standard errors. The resulting adjusted R^2 is slightly higher than 7% except for the baseline specification, in which it is de facto zero. These specifications, however, do not sufficiently control for all other changes that happen when parents shift jobs. Specifically, the potential effects coming from changes in income, work activities, etc. are all attributed to changes in technology exposure in the basic specifications. In contrast, the more comprehensive specifications in columns 4-5 in Table 4 control for income and occupational attributes that change with job shifts. These specifications suggest an average treatment effect of approximately 2%-points, which is significant at the 5% level using our preferred

[3] Additional tables are available upon request.

Table 4: Main regression results

	Regression #1	Regression #2	Regression #3	Regression #4	Regression #5
Treatment	0.0118*	0.0129*	0.0111	0.0175*	0.0216**
	(0.007)	(0.0068)	(0.0068)	(0.0092)	(0.0103)
Sex (male)		0.0696***	0.0696***	0.0696***	0.0695***
		(0.0016)	(0.0016)	(0.0016)	(0.0016)
Origin (non-Danish)		−0.0174	−0.0065	0.0047	0.0047
		(0.0307)	(0.031)	(0.0328)	(0.0329)
Citizenship (non-Danish)		−0.2618***	−0.2616***	−0.2623***	−0.2625***
		(0.0079)	(0.0079)	(0.008)	(0.008)
Experience			0.0558***	0.0042	0.0039
			(0.0138)	(0.0301)	(0.0301)
Health expenses			−0.0	−0.0	−0.0
			(0.0)	(0.0)	(0.0)
Hospital admissions			−0.0046	−0.0075	−0.0075
			(0.0079)	(0.0082)	(0.0081)
Sickness benefits			−0.0	0.0	0.0
			(0.0)	(0.0)	(0.0)
Income				0.0**	0.0**
				(0.0)	(0.0)
PCA1				−0.001	−0.0011
				(0.001)	(0.001)
PCA2				−0.0011	−0.0013*
				(0.0008)	(0.0008)
PCA3				−0.0034**	−0.0036**
				(0.0014)	(0.0014)
PCA4					0.0013
					(0.0015)
PCA5					0.0018
					(0.0014)
Adj. R^2	0.000	0.073	0.074	0.107	0.107
SSE	778.4	721.2	720.8	695.3	695.3
SSR	0.1	57.2	57.6	83.1	83.1
SST	778.4	778.4	778.4	778.4	778.4
N	139,183	139,183	139,183	139,183	139,183

Notes: This table shows the results from regressing the (predicted) income rank at ages 45 to 49 (the long-term outcome) on childhood exposure to technology (the treatment exposure) and additional control variables using five specifications that each represent a column. Column (1) represents the baseline specification with no controls, whereas column (5) represents the full specification with demographic variables, labor-market experience, health proxies, income, and occupational attributes. Columns (2)-(4) represent the specifications in between the baseline and full specification as detailed in the index column. Coefficients estimates are shown in the top of each cell alongside with standard errors in parentheses in the bottom. Superscripts ***, **, and * indicate statistical significance based on a (two-sided) t-test using cluster-robust standard errors at the family level at significance levels 1%, 5%, and 10%, respectively.

specifications with five principal components that represent the rich occupational data from O*NET of 244 occupational attributes. The adjusted R^2 is also higher at almost 11%. Adding further principal components does not change the estimated effects significantly. Altogether, Table 4 suggests that increasing the childhood exposure to technology at ages 5-15 by one standard error through parents shifting to jobs that involve a larger degree of interacting with computers will lead to the children rising approximately 2%-points in the income distribution when they reach their late 40s.

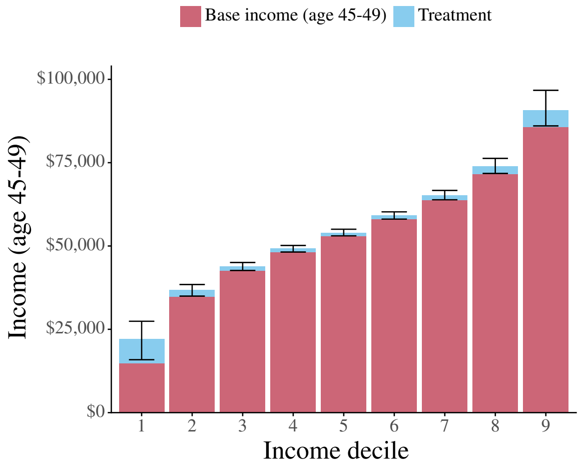


Figure 8: Effects of the treatment on the long-term outcome

Notes: This figure shows the estimated monetary effect on income in 2015 USD of childhood exposure to technology by income decile.

Heterogeneous effects Our finding that increased exposure to technology can positively affect children's opportunities to climb the income ladder by 2%-points is in itself a very heterogeneous effect because the monetary effect varies greatly with the income rank. To assess the effect heterogeneity, we compute the average monetary effect by decile of the income distribution and show the results in Figure 8. Note that we leave out the highest income decile because the effects are extremely large and distort the illustration, which follows from the fact that income in our sample follows a log-normal distribution. We provide the same figure including the tenth decile in Figure A.1 in Appendix A.

Figure 8 shows that the largest effects are to be expected at the tails of the income distribution, i.e., for the low- and high-earners. This makes sense if income follows a non-uniform distribution with most of the mass at the center, which is the case for our sample. Specifically, a 2%-points increase in income rank at the bottom of the income distribution (the first decile) translates to at 48% increase in income, which then corresponds to approximately $7,500 per year. Considering the median income rank, the estimated effect translates into slightly less than 2% increase in income, corresponding to around $1,000 extra per year. For individuals at the ninth decile, the increase in income is estimated to be just shy of 6%, which would be extra $5,200 per year. At the very top of the income distribution (not shown in Figure 8 but in Figure A.1), the increase in income is estimated to approximate 32%, which would bring in additionally $49,600 a year.

V. Conclusion

Technology affects every aspect of our economies and social welfare systems. Evidence suggests that the adoption of technologies currently happens at a faster pace that the creation of new tasks, and thereby the displacement effects dominates the reinstatement effects (Acemoglu and Restrepo, 2019a,b,c, 2020). One consequence is rising inequality (Acemoglu and Restrepo, 2022). However, to the best of our knowledge, there are no studies focusing on the impact of technology between generations. Specifically, one may hypothesize that parent's exposure to technology can act as a social leveler, enabling children from disadvantaged backgrounds to overcome their circumstances. This could materialize if parents who are highly exposed to technology pass on their interests and knowledge to their children who then develop high-demand skills (e.g., programming) or get employed in the high-paying technology sector. But lack of high-quality data on the subject (Frank et al., 2019) and the lack of sufficiently long panels of technology exposure and parent-child pairs create large barriers to study the effects of childhood exposure to technology on social mobility in terms of rising income rank. The topic, however, remains of the highest importance to policy-makers as a tool to develop the educational system of tomorrow.

This paper studies the impact of childhood exposure to technology at ages 5-15 on the long-term income rank at age 45-49. We exploit the Danish micro data that contain occupational codes and merge those with the detailed occupational data from O*NET. As exposure to technology, we use the degree to which using computers (hardware and software) is required to perform an occupation, which is most pronounced for computer systems engineers, computer research scientists, computer programmers, and akin. We adopt the multiple surrogates index framework by Athey et al. (2016) to overcome the challenge of having sufficiently long panels of parent-child pairs by modeling the effect of childhood exposure on adulthood income through intermediate surrogates, such as income, education, and labor-market experience at ages 25-29. Further, we extend the framework to allow for continuous treatments. Identifying the causal effects of technology, we follow Chetty and Hendren (2018a) in exploiting the variation in the age of children when their parents encounter shifting technology exposures, which then removes unobserved family heterogeneity. Further, as much more than exposure to technology changes with job shifts, we control for the financial situation, health conditions, and information from more than 240 occupational attributes.

Our findings indicate the individuals may benefits greatly in their adulthood from childhood exposure to technology. In fact, we estimate that a one standard error increase in the exposure to technology rises the income rank of the individual by 2%-points. The immediate policy recommendation that follows is to increase exposure to technology by expanding the computer-based activities in the educational curriculum, which would help children to acquire technology skills.

References

Acemoglu, D. and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, Elsevier, 1043–1171.

Acemoglu, D. and P. Restrepo (2019a): "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 33, 3–30.

——— (2019b): *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press, chap. Artificial Intelligence, Automation, and Work, 197–236.

——— (2019c): "The wrong kind of AI? Artificial intelligence and the future of labour demand," *Cambridge Journal of Regions, Economy and Society*, 13, 25–35.

——— (2020): "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 128, 2188–2244.

——— (2022): "Tasks, Automation, and the Rise in US Wage Inequality," *Econometrica (Forthcoming)*.

Agrawal, A., J. S. Gans, and A. Goldfarb (2019): "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction," *Journal of Economic Perspectives*, 33, 31–50.

Athey, S., R. Chetty, G. Imbens, and H. Kang (2016): "Estimating Treatment Effects using Multiple Surrogates: The Role of the Surrogate Score and the Surrogate Index," *arXiv preprint arXiv:1603.09326v3*.

Autor, D. H. (2015): "Why Are There Still So Many Jobs? The History and Future of Workplace Automation," *Journal of Economic Perspectives*, 29, 3–30.

Autor, D. H. and D. Dorn (2013): "The growth of low-skill service jobs and the polarization of the U.S. labor market," *American Economic Review*, 103, 1553–97.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006): "The polarization of the U.S. labor market," *American Economic Review*, 96, 189–194.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration*," *The Quarterly Journal of Economics*, 118, 1279–1333.

Becker, G. (1974): "A Theory of Social Interactions," *Journal of Political Economy*, 82, 1063–93.

Becker, G. and N. Tomes (1976): "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84, S143–62.

Becker, G. S. and N. Tomes (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 1153–1189.

Begg, C. B. and D. H. Y. Leung (2000): "On the Use of Surrogate End Points in Randomized Trials," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 163, 15–28.

Blau, P. M. and O. D. Duncan (1967): *The American Occupational Structure*, New York, Wiley.

Boudon, R. (1974): *Education, Opportunity, and Social Inequality*, New York, Wiley.

Bound, J. (1991): "Self-reported versus objective measures of health in retirement models," *The Journal of Human Resources*, 26, 106.

Bourdieu, P. (1973): *Cultural Reproduction and Social Reproduction*, Knowledge, Education, and Cultural Change (London: Tavistock Publications), Edited by Richard Brown.

——— (1977): *Reproduction in Education, Society, Culture.*, Sage, Beverly Hills, CA.

——— (1984): *Distinction: A Social Critique of the Judgement of Taste*, Cambridge, MA.: Harvard University Press,.

Brynjolfsson, E. and A. McAfee (2012): *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*, Brynjolfsson and McAfee.

——— (2014): *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*, Norton, New York.

Chetty, R., D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang (2017): "The fading American dream: Trends in absolute income mobility since 1940," *Science*, 356, 398–406.

Chetty, R. and N. Hendren (2018a): "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *The Quarterly Journal of Economics*, 133, 1107–1162.

——— (2018b): "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," *The Quarterly Journal of Economics*, 133, 1163–1228.

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

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

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014b): "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *American Economic Review*, 104, 141–147.

Darcy, A. M., A. K. Louie, and L. W. Roberts (2016): "Machine learning and the profession of medicine," *Jama*, 315, 551–552.

Dunis, C. L., P. W. Middleton, A. Karathanasopolous, and K. Theofilatos (2016): *Artificial Intelligence in Financial Markets: Cutting Edge Applications for Risk Management, Portfolio Optimization and Economics*, Springer.

Frangakis, C. E. and D. B. Rubin (2002): "Principal Stratification in Causal Inference," *Biometrics*, 58, 21–29.

Frank, M. R., D. Autor, J. E. Bessen, E. Brynjolfsson, M. Cebrian, D. J. Deming, M. Feldman, M. Groh, J. Lobo, E. Moro, et al. (2019): "Toward understanding the impact of artificial intelligence on labor," *Proceedings of the National Academy of Sciences*, 116, 6531–6539.

Frey, C. B. and M. A. Osborne (2017): "The future of employment: How susceptible are jobs to computerisation?" *Technological Forecasting and Social Change*, 114, 254–280.

Gustman, A. L. and T. L. Steinmeier (2018): "The role of health in retirement," in *Research in Labor Economics*, Emerald Publishing Limited, 229–297.

Holland, P. W. (1986): "Statistics and Causal Inference," *Journal of the American Statistical Association*, 81, 945–960.

Leontief, W. (1952): "MACHINES AND MAN," *Scientific American*, 187, 150–164.

Mokyr, J., C. Vickers, and N. L. Ziebarth (2015): "The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different?" *Journal of Economic Perspectives*, 29, 31–50.

Prentice, R. L. (1989): "Surrogate endpoints in clinical trials: Definition and operational criteria," *Statistics in Medicine*, 8, 431–440.

Rubin, D. B. (1974): "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of Educational Psychology*, 66, 688.

Solon, G. (1992): "Intergenerational Income Mobility in the United States," *The American Economic Review*, 82, 393–408.

Stern, S. (1989): "Measuring the effect of disability on labor force participation," *The Journal of Human Resources*, 24, 361.

Zimmerman, D. J. (1992): "Regression Toward Mediocrity in Economic Stature," *The American Economic Review*, 82, 409–429.

A. Appendix

In this appendix, we provide additional details to those given in Section IV on the application. This appendix is meant as a supplement and is not self-contained without the main text.

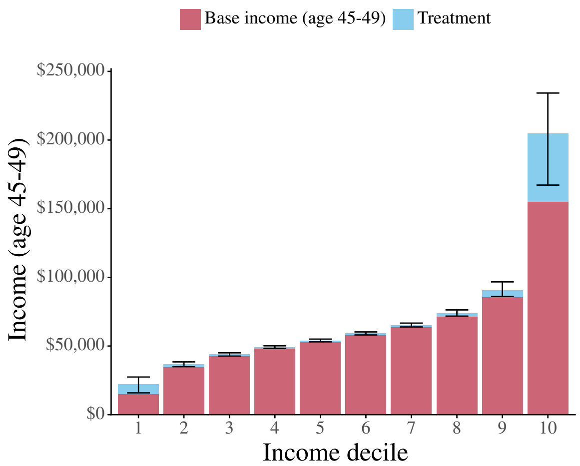


Figure A.1: Effects of the treatment on the long-term outcome

Notes: This figure shows the estimated monetary effect on income in 2015 USD of childhood exposure to technology by income decile.