Appendix

The Dynastic Benefits of Early Childhood Education: Participant Benefits and Family Spillovers

Contents

A1. Definition of Variables in Tables 2 and 3

Table A.1. Definition of Analysis Variables: Participants of PPP

Variable	***Description***
IQ	Participant's Stanford-Binet IQ test; baseline (age 3).
Socioeconomic Index	Participant's household socioeconomic index (based on parents' education, parents' employment, and rooms *per capita* in household); baseline (age 3).
Mother Does not Work	Indicator of participant's mother not currently working; baseline (age 3)
Mother's Birth Year	Participant's mother birth year; baseline (age 3).
Cognitive	Principal component factor of measures of general intelligence; midlife follow-up.
Non-Cognitive	Principal component factor of measures of personality; midlife follow-up.
High-School Graduate	Indicator of being a high-school graduate (does not include GED); midlife follow-up.
College Graduate	Indicator of being a four-year college graduate; midlife follow-up.
Physical Health	Principal component factor of diabetes indicator at age 40, home-care indicator at midlife, percentage of bedridden days during last year at midlife, and severe pain indicator at midlife; measure scales are reversed when appropriate.
Mental Health	Principal component factor of measures of depressed, anti-social, and community-engaged behavior; measure scales are reversed when appropriate; midlife follow-up.
Married	Fraction of years married between ages 20 and 40.
Labor Income	Average earnings from labor income between ages 20 and 40; 2017 USD.
Household Income	"Labor income" in addition to average spouse's labor income between ages 20 and 40; 2017 USD.
Total Days in Jail or Prison	Number of days in jail or prison up to the midlife follow-up.
Never Arrested	Never arrested indicator; midlife follow-up.
Any Children	Indicator of having any children; midlife follow-up.
$>$ 5 Children	Indicator of having more than five children; midlife follow-up.
Number of Children	Number of children; midlife follow-up.
Age at Onset	Age when first child born, among those who report having children; midlife follow-up.
Any Siblings	Indicator of having any children; midlife follow-up.
Number of Siblings	Number of siblings; midlife follow-up.
Up to 5 Eldest	Number of siblings among the five eldest, among those who report having siblings; midlife follow-up.
Fraction Older	Fraction of "up to 5 eldest siblings" who are older than participant; midlife follow-up.
Fraction Older; $<$ 5 Years Apart	Fraction of "up to 5 eldest siblings" who are older than participant but at most 5 years in age apart; midlife follow-up.

Note: This table describes the construction of the variables summarized in Table 2.

Table A.2. Definition of Analysis Variables: Children and Siblings of Participants of PPP

Variable	***Description***
Age	Age in years at the midlife follow-up.
High-School Graduate	Indicator of being a high-school graduate (does not include GED).
College Graduate	Indicator of being a four-year college graduate.
Employed	Indicator of being employed or self-employed, as opposed to unemployed, retired, or out of the labor force.
Employed or retired	Indicator of being employed or retired, as opposed to unemployed our out of the labor force.
Never Arrested	Never arrested indicator.
In Good Health	Indicator of health being good, very good, or excellent in a five-rating scale.
Not a Parent	Indicator of not having a child.
Never Divorced	Not divorced indicator.

Note: This table describes the construction of the variables summarized in Table 3.

A2. Details on the Treatment-Effect Estimators

A2.1 Overview

We use two basic estimators of the average treatment effect $\mathbb{E}\left[Y^1_{j,a} - Y^0_{j,a}\right]$ in Equation (2). The first is the average treatment-control difference (mean difference). For each outcome $j \in \mathcal{J}$, we pool the treatment and control groups at all observed ages to estimate

$$Y_{j,a} \; = \; \delta^0_{j,a} \; + \; \delta^1_{j,a} \, D \; + \; \varepsilon_{j,a} \;\; \text{for } a \in \mathcal{A}, \tag{A.1}$$

where $Y_{j,a}$ is the value of the observed outcome. In the framework of Quandt (1958), $Y_{j,a} := D \cdot Y^1_{j,a} + (1-D) \cdot Y^0_{j,a}$ where D indicates treatment status. $\delta^0_{j,a}$ and $\delta^1_{j,a}$ are age-specific coefficients and $\varepsilon_{j,a}$ is an error term. The coefficient $\delta^1_{j,a}$ is the mean difference for outcome $j \in \mathcal{J}$ at age $a \in \mathcal{A}$. Our interest is in the aggregate of the $\delta^1_{j,a}$ over $j \in \mathcal{J}$ and $a \in \mathcal{A}$, and not in any specific $\delta^1_{j,a}$.

Mean-difference estimators identify the average treatment effect, if treatment is randomized. To address randomization compromises described in Section 2, as well as the problem of attrition and missing data (i.e., item non-response), we use regression-adjusted mean differences (OLS) as our main estimator. While we do not know the exact form of randomization failure, we know that baseline (randomization) variables are only partially balanced across the treatment and control groups. We adjust the average treatment-control difference by including the baseline variables summarized in Table 2 as regressors in Equation (A.1). Doing so accounts for randomization compromises and allows missing-data patterns to vary across randomization covariates. Appendix A2 provides further details on our estimators.

As a check on our main OLS results, we also apply the robust methodology of Heckman and Karapakula (2021). It uses an augmented inverse-probability estimator (AIPW), which is a more general mean-difference adjustment than OLS. AIPW weights Equation (A.1) using the inverse probability of being treated and having non-missing data. The justification of AIPW is asymptotic. Its application to our sample may thus be inappropriate. Applying AIPW barely affects any of our results (see Appendix Table A.15).

A2.2 Technical Details

We provide a formal discussion on the estimators that we use in the paper, using the notation defined there. We define some additional notation to express our results formally. Let $a \in \mathcal{A} = \{\underline{a}, \underline{a}+1, \ldots, \overline{a}\}$ denote the ages of Perry participants between age $\underline{a}$ (treatment start) and $\overline{a}$ (the end of the life cycle) and let $\mathcal{P}$ index the unique identifiers of each of the 123 Perry participants. We partition $\mathcal{P}$ into index sets for the treatment and control groups, $\mathcal{P}_1$ and $\mathcal{P}_0$ respectively. Recall that we use the switching-regression notation of Quandt (1958, 1972) to denote outcome $j \in \mathcal{J}$ at age $\mathcal{A}$ as $Y_{j,a} = D \cdot Y^1_{j,a} + (1 - D)\, Y^0_{j,a}$. We drop the outcome index henceforth for brevity, and we introduce an individual index to make some of the calculations explicit.

Mean-Difference Estimator. The mean difference (**MD**) estimator assumes that missing data occur randomly. We define it as

$$\widehat{\Pi}_{\text{md}} := \sum_{a} \sum_{i \in \mathcal{P}_1} \frac{1}{N_{a,1,1}} \beta^{a-3} R_{i,a} Y^1_{i,a} - \sum_{a} \sum_{i \in \mathcal{P}_0} \frac{1}{N_{a,0,1}} \beta^{a-3} R_{i,a} Y^0_{i,a}, \tag{A.2}$$

where $R_{i,a} = 1$ indicates that the relevant variable is observed and $R_{i,a} = 0$ indicates that it is not. Note that Equation (A.2) is numerically equivalent to Equation (A.1) by the Frisch-Waugh-Lovell (FWL) theorem. We denote realizations of $R_{i,a}$ as r and realizations of D_i as d. $N_{a,d,r}$ is the number of observations in treatment status d observed at age a. $\widehat{\Pi}_{\text{md}}$ is a consistent estimator of the average treatment effect (ATE) under random assignment of treatment (**RA**). That is, $\{Y^1_{i,a}, Y^0_{i,a}\} \perp\!\!\!\perp D \;\; \forall i \in \mathcal{P}, a \in \mathcal{A}$. The **MD** estimator attaches equal weight to all of the components of averages of the outcomes across ages, and therefore corrects for general age-driven patterns of missing data over the life-cycle.[1] For the **MD**, we hence assume $\{Y^1_{i,a}, Y^0_{i,a}\} \perp\!\!\!\perp R_{i,a} \;\; \forall i \in \mathcal{P}, a \in \mathcal{A}$ (**MAR I**).

The **MD** is a baseline estimator. However, the randomization protocol of PPP only justifies conditional random assignment. We use an **OLS** estimator to account for compromises in the randomization protocol. We run a pooled regression of the outcome variable $Y_{i,a}$ on a vector of baseline variables, Z_i, and on age indicators interacted with treatment status. We denote the slope vector associated with Z_i by γ. We correct $Y_{i,a}$ for compromises in the randomization protocol and missingness patterns depending on Z_i by forming $Y_{i,a} - \widehat{\gamma}' Z_i$, and use this quantity instead of $Y_{i,a}$ in the formula for the **MD** estimator. This linear regression based correction removes an individual effect $\gamma' Z_i$ correlating with assignment and missingness patterns. The **OLS** estimator is consistent and unbiased for the ATE given conditional random assignment $\{Y^1_a, Y^0_a\} \perp\!\!\!\perp D \mid Z$ (**CRA**), missingness at random conditional on age and Z_i, $R_a \perp\!\!\!\perp \{Y^1_a, Y^0_a, D\} \mid Z$ (**MAR II**) and the specification assumption $\mathbb{E}[Y^d_a \mid R = 1, D = d, Z] = \alpha_{a,d} + \gamma' Z, \forall a \in \mathcal{A}$ ($\mathbf{S}_{\text{OLS}}$).

We also consider an *augmented inverse probability weighting* (**AIPW**) estimator. Compared to **OLS**, this estimator allows treatment assignment and missing data patterns to depend on Z_i in a more general fashion, by relaxing specification assumptions. We assume **CRA** and **MAR II** and make the specification assumption $\mathbb{E}[Y^d_a \mid R = 1, D = d, Z] = \alpha_{a,d} + \gamma'_{a,d} Z$ *or* $\mathbb{P}(R_a = 1, D = 1 \mid Z) = \Lambda\big([1, Z']\omega^R_a\big)\Lambda\big([1, Z']\omega^D_a \mid R_a = 1\big)$ ($\mathbf{S}_{\text{AIPW}}$). Here, Λ denotes the Logit function, $[1, Z']$ denotes the row vector of baseline covariates concatenated with 1, and $\omega^R_a, \omega^D_a \in \mathbb{R}^{\dim(Z)}$. The fact that only one of the two preceding equations needs to hold

[1] It considers the ages sampled in our data, not all possible ages that could be sampled (i.e., it assigns zero probability to ages not sampled in our data).

is also known as *double robustness*, an appealing property of the **AIPW** estimator.[2] We construct **AIPW** estimates as follows. We let $\widehat{Y}_{i,a}$ be an estimate of $\mathbb{E}[Y_{i,a} \mid Z_i, D_i, R_{i,a} = 1]$, the expected outcome of i at age a, conditional on treatment status, non-missing data, and Z_i. Additionally, the **AIPW** estimator uses estimates $\widehat{\phi}_i^d$ of $\phi_i^d := \mathbb{P}(D_i = d \mid Z_i)$ (i.e., the i-th participant's propensity of being in the treatment status d) and $\widehat{\lambda}_{i,a}^d$, an estimator of $\lambda_{i,a}^d := \mathbb{P}(R_{i,a} = 1 \mid Z_i, D_i = d)$, the propensity of having a non-missing outcome after fixing treatment status D_i to $d \in \{0, 1\}$. The estimator is

$$\widehat{\Pi}_{\text{aipw}} = \frac{1}{N_{\mathcal{P}}} \sum_{i \in \mathcal{P}} \sum_{a} \beta^{-3} \left(\widehat{\theta}_{i,a}^1 - \widehat{\theta}_{i,a}^0 \right), \tag{A.3}$$

where

$$\widehat{\theta}_{i,a}^d := \widehat{Y}_{i,a}^d + \frac{\mathbf{1}\{R_{i,a} = 1, D_i = d\}}{\widehat{\lambda}_{i,a}^d \widehat{\phi}_i^d} \left(Y_{i,a}^d - \widehat{Y}_{i,a}^d \right),$$

and where $\mathbf{1}(\cdot)$ is the indicator function. This **AIPW** estimator is doubly robust: either correct specification of (1) the propensity score models for $\widehat{\phi}_{i,a}^d$ and $\widehat{\lambda}_{i,a}^d$ or (2) the model for $\widehat{Y}_{i,a}^d$ for $d \in \{0, 1\}$ implies consistency. The imputation scheme of the **AIPW** estimator allows us to choose outcome-domain specifications of $\widehat{Y}_{i,a}^d$. In particular, we can model censored outcome variables explicitly.

Table A.3 summarizes the requirements for each estimator to be consistent. The proof of consistency of the **MD** estimator is straightforward. We provide proofs of consistency for **OLS** and **AIPW** next.

A2.3 Proof of Consistency of the OLS Estimator

First, make assumption $\mathbf{S}_{\text{OLS}}$. Then, we can write $Y_a^d = \alpha(a, d) + \gamma' Z + e(a, d)$. For fixed age and treatment status, $\alpha(a, d)$ is a constant, γ is a slope vector, and $e(a, d)$ is a stochastic, zero mean error. Because of **CRA** and **MAR II**, $\mathbb{E}[e(a, d) \mid R, D, Z] = 0$ holds. Using the

[2] Note, however, that the **AIPW** estimator is, in contrast to **MD** and **OLS**, not unbiased.

Table A.3. Assumptions Required For Consistency of Each Estimator

Assumption (for all $a \in \mathcal{A}$)		**Estimator** MD	OLS	AIPW
Missing-Data Assumptions				
MAR I	$R_a \perp\!\!\!\perp \{Y_a^1, Y_a^0, D\}$	×		
MAR II	$R_a \perp\!\!\!\perp \{Y_a^1, Y_a^0, D\} \mid Z$		×	
Treatment Assignment Assumptions				
RA	$(Y_a^1, Y_a^0) \perp\!\!\!\perp D$	×		
CRA	$(Y_a^1, Y_a^0) \perp\!\!\!\perp D \mid Z$		×	×
Specification Assumptions				
$\mathbf{S}_{\text{OLS}}$	$\mathbb{E}[Y_a^d \mid R = 1, D = d, Z] = \alpha_{a,d} + \gamma' Z$		×	
$\mathbf{S}_{\text{AIPW}}$	$\mathbb{E}[Y_a^d \mid R = 1, D = d, Z] = \alpha_{a,d} + \gamma'_{a,d} Z$ **or** $\mathbb{P}(R_a = 1, D = 1 \mid Z) = \Lambda\big([1, Z']\omega_a^R\big)\Lambda\big([1, Z']\omega_a^D \mid R_a = 1\big)$			×
Overlap Assumptions				
Overlap D	Common covariate support of treatment and control group. $1 - \varepsilon > \mathbb{E}[D \mid Z] > \varepsilon$ for some $\varepsilon > 0$, all Z.		×	×
Overlap R	Common covariate support of population with missing and non-missing observations. $\mathbb{E}[R \mid Z] > \varepsilon$ for some $\varepsilon > 0$, all Z.		×	×

switching regression framework $Y_a = DY_a^1 + (1-D)\,Y_a^0$ we rewrite the specification equation into the linear regression equation

$$Y_a = \alpha(a,0) + \Delta_a D + \gamma' Z + e(a),$$

where we let $\Delta_a := (\alpha(a,1) - \alpha(a,0))$ and $e(a) := e(a,0)D + e(a,0)(1-D)$, and $\mathbb{E}[e(a) \mid D, Z] = 0$ holds. We estimate γ using linear regression of the pooled sample of outcomes $(Y_{i,a})_{i\in\mathcal{P}, a\in\mathcal{A}}$ on Z_i, D_i and a full set of age dummies interacted with D and the intercept. We write

$$Y_{i,a} - \widehat{\gamma} Z_i = \begin{cases} o_p(1) + \alpha(a,0) + e(a)_i & \text{if } D_i = 0 \\ o_p(1) + \alpha(a,1) + e(a)_i & \text{if } D_i = 1, \end{cases} \tag{A.4}$$

where $e(a)_i$ may be heteroskedastic and correlated or clustered within households. Therefore, substituting $Y_{i,a} - \widehat{\gamma} Z$ for $Y_{i,a}$ in the **MD** formula yields (dropping discount rates for brevity and with $N_{a,d,r}$ defined as before)

$$\widehat{\Pi}_{\text{ols}} = \sum_a \sum_{i\in\mathcal{P}_1} \frac{1}{N_{a,1,1}} R_{i,a}(Y_{i,a}^1 - \widehat{\gamma}' Z_i) - \sum_a \sum_{i\in\mathcal{P}_0} \frac{1}{N_{a,0,1}} R_{i,a}(Y_{i,a}^0 - \widehat{\gamma}' Z_i), \tag{A.5}$$

and consistency follows from

$$\sum_{i\in\mathcal{P}_d} \frac{1}{N_{a,d,1}} R_{i,a}(Y_{i,a}^d - \widehat{\gamma}' Z_i) = o_p(1) + \mathbb{E}\left[\alpha(a,d) + e(a)_i \mid D = d, R_a = 1\right] + \left[\sum_{i\in\mathcal{P}_d} \frac{R_{i,a}}{N_{a,d,1}} o_p(1)\right].$$

The last term converges quickly as it is of stochastic order $o_p(\sqrt{N_{a,d,1}}^{-1})$ and can therefore be ignored in variance calculations. We then use $\mathbb{E}\left[\alpha(a,d) \mid D = d, R_a = 1\right] = \alpha(a,d)$ and form $\alpha(a,1) - \alpha(a,0) = \left(\mathbb{E}[Y_a^1] - \gamma'\mathbb{E}[Z]\right) - \left(\mathbb{E}[Y_a^0] - \gamma'\mathbb{E}[Z]\right) = \mathbb{E}[Y_a^1 - Y_a^0]$. Therefore, $\widehat{\Pi}_{\text{ols}}$ sums consistent estimates of age-wise treatment effects, which implies consistency of $\widehat{\Pi}_{\text{ols}}$. □

A2.4 Proof of Double Robustness of the AIPW Estimator

In the following discussion, we ignore D_i for simplicity, given that we assume that D_i is randomly assigned conditional on Z_i.

Let $\widehat{\lambda}_{i,n} = p(Z_i) + o_p(1)$ and $\widehat{Y}_{i,n} = m(Z_i) + o_p(1)$ for some functions $Z_i \mapsto p(Z_i) \in (\varepsilon, 1]$, some $\varepsilon > 0$ ($p(\cdot)$ is bounded away from zero) and $Z_i \mapsto m(Z_i) \in \mathbb{R}$. Define $\widehat{\theta}_n = n^{-1} \sum_{i=1}^{n} \widehat{\theta}_{i,n}$, where $\widehat{\theta}_{i,n} = \widehat{Y}_{i,n} + (R_i/\widehat{\lambda}_{i,n})\,(Y_{i,n} - \widehat{Y}_{i,n})$. The next assumption states that either the model for $\widehat{\lambda}_{i,n}$ or the model for $\widehat{Y}_{i,n}$ or both are correctly specified.

Specification Assumption. $p(Z_i) := \mathbb{E}[R_i \mid Z_i]$ or $m(Z_i) := \mathbb{E}[Y_i \mid Z_i]$ (or both) hold. (This assumption is a general version of $\mathbf{S}_{\text{AIPW}}$, leaving the concrete specification of m and p open.)

Proposition. If **MAR II**, **CRA**, $\mathbf{S}_{\text{AIPW}}$ hold, then $\widehat{\theta}_n = \mathbb{E}[Y_i] + o_p(1)$.

Proof. Note that $\widehat{\theta}_n = [n^{-1} \sum_{i=1}^{n} m(Z_i) + (R_i/p(Z_i))\,(Y_i - m(Z_i))] + o_p(1)$, and so

$$
\begin{aligned}
\widehat{\theta}_n &= \mathbb{E}[m(Z) + \tfrac{R}{p(Z)}(Y - m(Z))] + o_p(1) \qquad \text{(A.6)}\\
&= \mathbb{E}[\mathbb{E}[m(Z) + \tfrac{R}{p(Z)}(Y - m(Z)) \mid Z]] + o_p(1)\\
&= \mathbb{E}[m(Z) + \tfrac{1}{p(Z)}\mathbb{E}[R\,Y \mid Z] - \tfrac{\mathbb{E}[R \mid Z]}{p(Z)} m(Z)] + o_p(1)\\
&\overset{(\mathbf{MAR}\,II)}{=} \mathbb{E}[m(Z) + \tfrac{\mathbb{E}[R \mid Z]}{p(Z)}\mathbb{E}[Y \mid Z] - \tfrac{\mathbb{E}[R \mid Z]}{p(Z)} m(Z)] + o_p(1)\\
&= \mathbb{E}[m(Z) + \tfrac{\mathbb{E}[R \mid Z]}{p(Z)}(\mathbb{E}[Y \mid Z] - m(Z))] + o_p(1).
\end{aligned}
$$

If the propensity score model is correctly specified, i.e., $p(Z) := \mathbb{E}[R \mid Z]$, then

$$\widehat{\theta}_n = \mathbb{E}[m(Z) + \tfrac{p(Z)}{p(Z)}(\mathbb{E}[Y \mid Z] - m(Z))] \tag{A.7}$$
$$+ o_p(1) = \mathbb{E}[\mathbb{E}[Y \mid Z]] + o_p(1) = \mathbb{E}[Y] + o_p(1).$$

If the regression model is correctly specified, i.e., $m(Z) := \mathbb{E}[Y \mid Z]$, then

$$\widehat{\theta}_n = \mathbb{E}[\mathbb{E}[Y \mid Z] + \frac{\mathbb{E}[R \mid Z]}{p(Z)}(\mathbb{E}[Y \mid Z] - \mathbb{E}[Y \mid Z])] \tag{A.8}$$
$$+ o_p(1) = \mathbb{E}[\mathbb{E}[Y \mid Z]] + o_p(1) = \mathbb{E}[Y] + o_p(1).$$

Under our specification assumption, $p(Z) := \mathbb{E}[R \mid Z]$ or $m(Z) := \mathbb{E}[Y \mid Z]$. Therefore, $\widehat{\theta}_n = \mathbb{E}[Y] + o_p(1)$. □

Note that our proposition can be applied without loss of generality to the crime **AIPW** estimator described in the main text. To see why, index all variables (except for Z_i) with a fixed superscript $d \in \{0, 1\}$ in the assumptions and the theorems above.[3] Similar changes occur to the notation in the other assumptions and results.

Then, the modified Theorem 1 implies that $\widehat{\theta}_n^d = \mathbb{E}[Y_i^d] + o_p(1)$ under our assumptions. Therefore, $\widehat{\theta}_n^1 - \widehat{\theta}_n^0 = \mathbb{E}[Y_i^1 - Y_i^0] + o_p(1)$, proving that the crime **AIPW** estimator of the treatment effect is doubly robust to certain forms of misspecification.

A2.5 Estimation of AIPW Conditional Expectations and Probabilities

Recall that $\widehat{Y}_{i,a}^d$ is an estimate of $\mathbb{E}[Y_{i,a} \mid Z_i, D_i = d, R_{i,a} = 1]$ for $d \in \{0, 1\}$ for individual $i \in \mathcal{P}$ at age a. At some ages, the variation in the dependent variable, $Y_{i,a}$, may be too little to reliably estimate the desired conditional expectation function. If we used observations at

[3]These assumptions follow because, for the original **AIPW** estimators of the treatment effect, we assume that $(C_i^d, Y_i^d) \perp\!\!\!\perp D_i \mid Z_i$ and $(C_i, Y_i) \perp\!\!\!\perp R_i \mid D_i, Z_i$. In other words, $(C_i^d, Y_i^d) \perp\!\!\!\perp \mathbf{1}(D_i = d) \mid Z_i$, while $(C_i^d, Y_i^d) \perp\!\!\!\perp R_i^d \mid Z_i$.

age a only, an estimation may fail in bootstrap samples. Estimating the desired conditional expectation pooling all ages may introduce bias since that assumes the expected outcome, conditional on covariates Z_i (and with fixed treatment and no missing-data status) to be age invariant. As a middle ground, we estimate weighted regressions. The weights are obtained from a normal density kernel. Given age a, the weights attached to adjoining observations at ages $a \pm k, k = 0, 1, 2, ...$ are $\phi\left(\frac{|k|}{b_l}\right)$, where b_l is a bandwidth parameter specific for each domain. Estimation results are insensitive to the choice of the bandwidth. We use a bandwidth of 1 for crime and a bandwidth of 4 for income. We use no weighting in health and education.

We define $\widehat{Y}^d_{i,a}$ as the prediction of $Y^d_{i,a}$ derived from some regression model of $Y^d_{i,a}$ on Z_i in the (alive) population $\{i \in \mathcal{P} : D_i = d, R_{i,a} = 1\}$. The regression model $Y^d_{i,a}$ is specified according to each domain that we consider.

We account for the PPP participants who have died at some age $a \in \mathcal{A}$ by excluding them from the estimation sample for $\widehat{Y}^d_{i,a}$ when predicting outcomes of their living peers. Among participants, mortality was high (12%), with 10% mortality in the treatment and 14% mortality in the control group. Generally, the outcome $Y_{i,a}$ is nil for deceased individuals. Formally, let $\mathfrak{D}_{i,a}$ be an indicator for whether individual i at age a is deceased. In all preceding and following considerations, $\mathfrak{D}_{i,a}$ can be interpreted as a variable in $Z_{i,a}$ (adding an age index to Z_i), without loss of generality. All our models for missing-data probabilities build on a penalized Logit specification in Greenland and Mansournia (2015) mixed with unit probability for deceased individuals.[4] For instance, consider the missing-data model for

[4] Using the penalized Logit regression in Greenland and Mansournia (2015) guarantees that the Logit model remains estimable even if, for some ages, there is little variation in covariates and outcomes. Penalization is derived from imposing a $\log[F(1,1)]$ prior on each coefficient in the Logit model. The advantages of this prior are 1) finite estimates on all coefficients even with perfect separation, 2) it constitutes a direct bias reduction method, and 3) it is easy to implement via a simple data augmentation.

$R_{i,a}$ for felonies. We estimate

$$\widehat{\lambda}_{i,a}^{d} = (1 - \mathfrak{D}_{i,a}) + \mathfrak{D}_{i,a}\widehat{\Lambda}(R_{i,a} \mid Z_i, \mathfrak{D}_{i,a} = 0, D_i = d)$$

where $\widehat{\Lambda}(R_{i,a} \mid Z_i, \mathfrak{D}_{i,a} = 0, D_i = d)$ is a penalized Logit specification with covariates Z_i, estimated in the sample of non-deceased individuals with treatment status d.

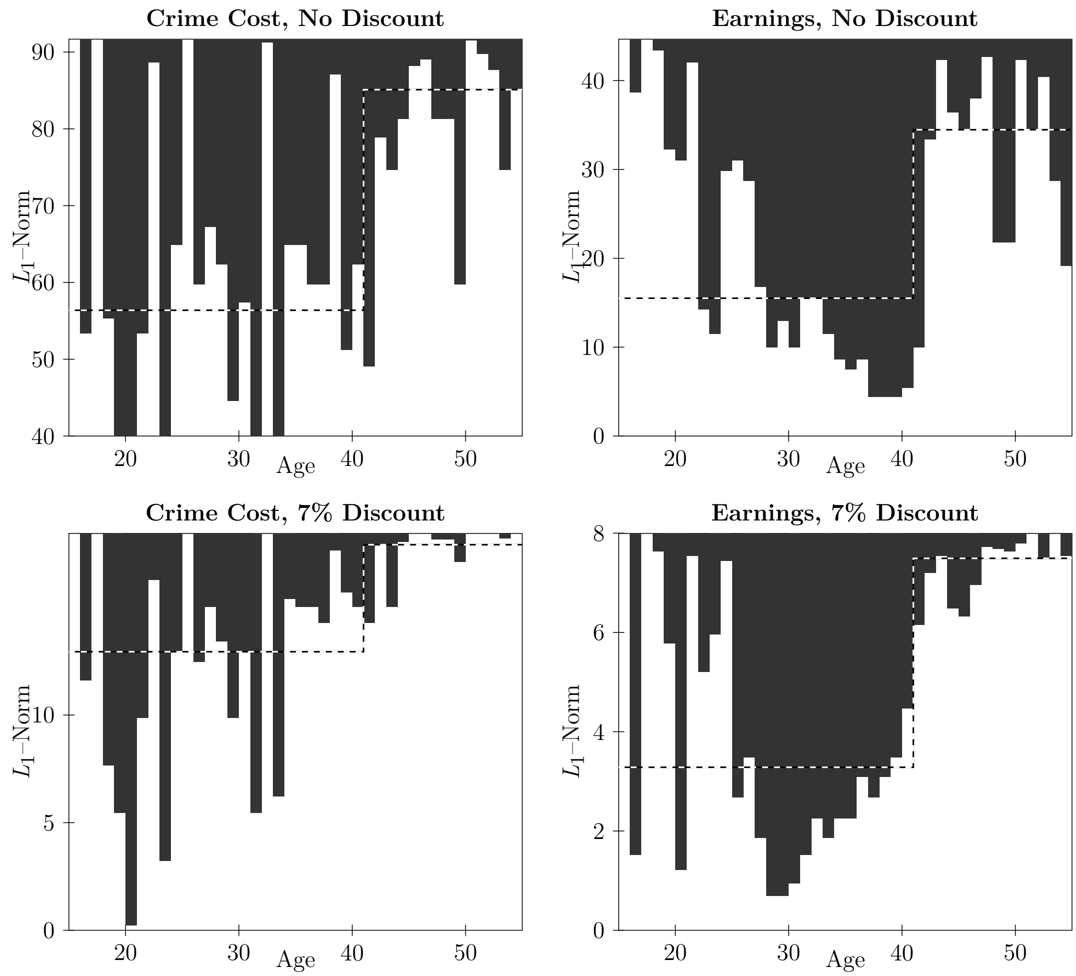
A2.6 Selecting Treatment-Effect Age Ranges Using LASSO

Our preferred set of results exclude treatment effects on labor income and crime after age 40. After age 40, the treatment-effect estimates are minimal in magnitude and do not differ from 0 statistically. Including them increases the variance of our life-cycle estimates. We justify the age-40 cutoff using a least-absolute shrinkage and selection operator (LASSO) generalization of Equation (A.1). We vary the penalty parameter λ away from the **OLS** solution ($\lambda = 0$), until the LASSO selects no age-wise treatment-effect components. We penalize the estimator for including coefficients $\delta_{j,a}^{1} \neq 0$ for a large number of ages $a \in \mathcal{A}$, but apply no penalty to the age intercepts $\delta_{j,a}^{0} \neq 0$. Figure A.1 plots the order in which LASSO picks up the explanatory variables and associated coefficients $\delta_{j,a}^{1} \neq 0$. The age-40 cutoff is clear.

A3. Details on Inference Procedures

Our bootstrap procedures are stratified to keep the size of the treatment and control groups roughly equal across resamplings. They account for sampling variation in all estimation stages required by our estimators (i.e., we form the empirical bootstrap distributions of our estimators by computing all estimation stages in each bootstrap draw). All of our inference on the participants is clustered at the household level, defining households as individual-sibling clusters. We modify the clustering of our inference on the spillover benefits as explained

Figure A.1. LASSO Coefficients L_1-Norm vs Included Ages



Notes: The x-axis shows the ages at which treatment effects were measured. The y-axis displays the L_1–norm of the standardized coefficient vector in a LASSO regression when varying the penalty parameter (λ). Dark sections indicate inclusion of age-wise treatment effects if the LASSO produces a given L_1–norm. Higher L_1 values correspond to lower penalties and a LASSO more similar to the **OLS** solution. The dashed line displays median L_1–norm at variable inclusion, applicable to age ranges up to 40 and past 40, respectively.

in Section 4. Appendix A6 compares our main inference to inference based on standard asymptotic methods and alternative bootstrap procedures.

Our main inference is based on bias-corrected accelerated confidence intervals (BCAs). In the paper, we present 90% symmetric confidence intervals (i.e., we leave 5% of mass to the left of the lower bound of the interval and 5% mass to the right of the upper bound of the interval). However, our inference is based on one-sided tests. We thus invert non-symmetric confidence intervals to obtain BCA p-values. Throughout the appendix, we also provide various bootstrap standard errors and p-values. Appendix A.15 shows that the inference based on BCAs or based on analytic, simple bootstrap, trimmed bootstrap, or studentized bootstrap p-values is very similar across estimates obtained using our different estimators.

All of the bootstrap inference procedures account for sampling variation in all estimation stages (i.e., we form the bootstrap distributions of our estimates by computing all of the stages required by our estimators in each draw). We also account for simulation or forecasting error in our health predictions as explained below. We provide details on each of our inference procedures next.

A3.1 Bias-Corrected Accelerated Bootstrap Confidence Intervals

We construct the BCAs as indicated in Efron (1987). Hansen (Chapter 10.18, 2021) provides additional discussion. The BCAs are fully non-parametric, but they are computationally intensive partly because they require the estimation of two parameters. First, a measure of small-sample median bias. This is the standard correction parameter in bias-corrected inference procedures. Second, a measure of skewness of the distribution of the parameter of interest. This feature is specific to the BCAs and it enables accounting for skewness in outcome distributions. The advantage of the inference provided by the BCAs with respect to other methods like bias-corrected percentile-t (or studentized) bootstrap confidence intervals is that it explicitly accounts for skewness as a small-sample anomaly. Inspection of bootstrap distributions indicates that skewness is present in the present value of some of our outcomes, especially when using the **AIPW** estimator. We choose BCAs for our main inference because they allow us to deal explicitly with small-sample size and skewness. Furthermore, BCAs and percentile-t bootstrap inference, which we discuss below, provide an asymptotic refinement compared to traditional inference methods. This refinement makes them the most accurate inference methods we employ.

Let $(\boldsymbol{Y}, \boldsymbol{X})_i$ denote the outcomes and covariates of PPP participant $i \in \mathcal{I}$ at all of the observed ages, where $\mathcal{I}$ is the index set for the PPP participants. The covariates include all of their information in each of their observed ages (i.e., baseline characteristics, treatment status, missing-data indicators). We partition individuals into their households (i.e., form sibling tuples). We let $\mathcal{P}$ index households and $(\boldsymbol{Y}, \boldsymbol{X})_h$ denote the outcomes and covariates of PPP participant household $h \in \mathcal{P}$. Our step-by-step bootstrap procedure to form the empirical bootstrap distribution of estimator θ is the following:

1. Draw $b = 1, \ldots, B$ bootstrap samples $(\boldsymbol{Y}, \boldsymbol{X})_h^b$ with replacement with the restriction of keeping the size of the treatment and control groups constant across draws. In this and all bootstrap procedures in the paper we set B to 1,000.

2. For each $b = 1, \ldots, B$, decompose $(\boldsymbol{Y}, \boldsymbol{X})_h^b$ into the individual information of the household participants. This enables forming $(\boldsymbol{Y}, \boldsymbol{X})_i^b$ for $i \in \mathcal{I}_b$ where $\mathcal{I}_b$ indexes individuals in bootstrap sample b.

3. For each $b = 1, \ldots, B$, estimate all preliminary stages (e.g., weighting scheme for **AIPW** estimates).

4. For each $b = 1, \ldots, B$, estimate main parameter. Denote it by $\widehat{\theta}^b$.

5. Form empirical bootstrap distribution $\widehat{\theta}^1, \ldots, \widehat{\theta}^B$.

Health Outcomes: A Special Case. Our health outcomes are based on modeling and simulation as explained in Appendix A5.4. We expand the bootstrap-sampling procedure to account for forecasting error in the simulated outcomes. Our step-by-step bootstrap procedure forms the empirical bootstrap distribution of estimator θ which contains a health outcome is the following.

1. Recall that for each $i \in \mathcal{I}$ at age $a \in \mathcal{A}$ we have 1,000 simulated health outcomes. The point estimate for any health outcome is the average across the 1,000 simulated outcomes. We form the individual and age-specific vector of residuals for each health outcome by forming the deviation of simulated outcome $s = 1, \ldots, S$ from the outcome's point estimate. Let $\mathcal{E}_{i,a}$ denote the vector storing these residuals for $i \in \mathcal{I}$ at age $a \in \mathcal{A}$. This vector stores an individual and age specific empirical distribution of outcome forecasting error.

2. Draw $b = 1, \ldots, B$ bootstrap samples $(\boldsymbol{Y}, \boldsymbol{X})_h^b$ with replacement with the restriction of keeping the size of the treatment and control groups constant across draws.

3. For each $b = 1, \ldots, B$, decompose $(\boldsymbol{Y}, \boldsymbol{X})_h^b$ into the individual information of the household participants. This enables forming $(\boldsymbol{Y}, \boldsymbol{X})_i^b$ for $i \in \mathcal{I}_b$ where $\mathcal{I}_b$ indexes individuals in bootstrap sample b.

4. For $i \in \mathcal{I}_b$ at age $a \in \mathcal{A}$, draw a residual from $\mathcal{E}_{i,a}$ for each health outcome and add it to its point estimate.

5. For each $b = 1, \ldots, B$, estimate all preliminary stages (e.g., weighting scheme for **AIPW** estimates).

6. For each $b = 1, \ldots, B$, estimate main parameter. Denote it by $\widehat{\theta}^b$.

7. Form empirical bootstrap distribution $\widehat{\theta}^1, \ldots, \widehat{\theta}^B$.

A3.2 Simple Bootstrap Standard Errors and p-values

The simple bootstrap standard errors are the standard deviations of the empirical bootstrap distributions (which we construct as explained above). We calculate the p-values associated with simple standard errors using t-statistics.

A3.3 Trimmed Bootstrap Standard Errors and p-values

The trimmed bootstrap standard errors are the standard deviations of the trimmed empirical bootstrap distributions. After obtaining the empirical distributions as explained above, we trim the top 1.0% and bottom 1.0% before computing the standard errors. Chapter 10 of Hansen (2021) notes that the trimmed bootstrap leads to more reliable standard errors (given that the trimming parameter vanishes as sample size tends to infinity). Note that we do not apply trimming to any other estimator. We calculate the p-values associated with trimmed standard errors using t-statistics.

A3.4 Percentile-t or Studentized Bootstrap p-values

An alternative non-parametric p-value is based on the studentized empirical bootstrap distributions as in Heckman and Karapakula (2021), also known as percentile-t bootstrap. Like the BCAs, the percentile-t method provides an asymptotic refinement. However, our use of analytic standard errors for studentization makes this method not fully non-parametric. Its p-values are calculated from the empirical bootstrap distribution of the t-statistic associated with the null hypothesis to be tested. We calculate studentized p-values as the fraction of draws in which the sampled statistic is more extreme (with respect to the null hypothesis) than the statistic in the original sample.

A3.5 Analytic Standard Errors by Outcome

For the **OLS** and **AIPW** estimators, the sampling variation from correctly specified preliminary estimation stages does not matter asymptotically (there are no preliminary estimation stages in **MD**). Hence it suffices to consider p-values calculated from final-stage regressions.[5] We calculate analytic standard errors allowing for general heteroskedasticity and arbitrary correlation within households. We calculate asymptotic analytic p-values based on the corresponding analytic standard errors. We use clustered standard errors robust to heteroskedasticity as in Liang and Zeger (1986), with a simple multiplicative degrees-of-freedom bias adjustment. Preliminary calculations indicate that alternative robust small-sample methods as those in Bell and McCaffrey (2002) and Imbens and Kolesar (2016) yield virtually identical results. Such methods include improved bias adjustments and adjusted reference distributions for confidence intervals and p-values.

Let Q be the design matrix of the regression in Equation (A.1). We estimate the variance of the estimated coefficients by computing

$$\widehat{\mathbb{V}}_{\text{md}} = (QQ')^{-1}\left(\sum_{h\in\mathcal{H}} Q'_h \widehat{\varepsilon}\widehat{\varepsilon}' Q_h\right)(QQ')^{-1}.$$

This is the cluster-robust standard error procedure of Liang and Zeger (1986). $\mathcal{H}$ is the set of all households in the study, and $(Q_h, \widehat{\varepsilon}_h)$ are the portions of the design matrix and residual vector that correspond to household h. The corresponding variance estimates $\widehat{\mathbb{V}}_{\text{ols}}$ for the **OLS** adjusted estimator and $\widehat{\mathbb{V}}_{\text{aipw}}$ for the **AIPW** are constructed analogously.

There are various bias-adjustment methods for clustered standard errors. We performed preliminary exercises using the methods in Bell and McCaffrey (2002) and Imbens and Kolesar (2016), which did not produce notable differences in our baseline standard-error estimates. Imbens and Kolesar (2016) suggest using 1) a bias improved estimator of the variance-

[5] The variance of the **AIPW** estimator is not doubly robust. We ignore this potential issue for analytic standard errors to keep things simple and highlight additional results, presented in Appendix A6, with respect to bootstrapped p-values, which are asymptotically correct even under misspecification of one of the first-stage models.

covariance matrix, $\widehat{\mathbb{V}}_{BM}$, as in Bell and McCaffrey (2002), and 2) comparing t-statistics of the kth coefficient relevant parameter to a t-distribution with K degrees of freedom, where K is calculated so that the distribution of the squared t-statistic of the kth estimated coefficient fits the first two moments of a $\chi^2(K)$ distribution. Neither approach makes a difference in our case. One minor degrees-of-freedom adjustment that we make is multiplying estimates $\widehat{\mathbb{V}}_{md}$, $\widehat{\mathbb{V}}_{ols}$ and $\widehat{\mathbb{V}}_{aipw}$ by the factor $\dfrac{N-1}{N-L}\dfrac{S}{S-1}$, where L denotes the number of estimated parameters in the model and S the number of clusters (households). We calculate p-values using the quantiles of the standard normal distribution.

A3.6 Analytic Standard Errors of Aggregate Estimator

To simplify our algebra, we note that the **MD** estimator can be written as the weighted sum of observations, $\widehat{\Pi}_{\text{md}} = W'\mathbf{Y}$, where

$$W = (w_i)_{i=1}^{|\mathcal{P}|},\ w_i = \frac{(2D_i - 1)R_{i,a}}{D_i N_{a,1,1} + (1 - D_i)N_{a,0,1}}$$

and $N_{a,d,r} = |\{i \in \mathcal{P} : R_{i,a} = r, D_i = d\}|$. The same weights are used to construct $\widehat{\Pi}_{\text{ols}} = W'(\mathbf{Y} - \widehat{\gamma}'\mathbf{Z})$, and weights for the **AIPW** estimator are given by $w_i = 2D_i/|\mathcal{P}|$, where $\widehat{\Pi}_{\text{aipw}} = W'(\widehat{\boldsymbol{\theta}}^1 - \widehat{\boldsymbol{\theta}}^0)$.

The variance of $\widehat{\Pi}_\Sigma$, $\mathbb{V}_\Sigma$, can be broken down into the variances of the estimators of individual domains j, $\mathbb{V}_j$, and the covariances between them, $\mathbb{V}_{j,\widetilde{j}}, j, \widetilde{j} \in \mathcal{J}$. We can write any of our estimators constructed for domain j as a weighted sum of $U_{i,a}$, where $U_{i,a}$ is either the observed outcome (**MD**), the regression adjusted observed outcome (**OLS**), or the imputed individual treatment effect (**AIPW**) for individual i at age a, depending on the estimator. Therefore, $\mathbb{V}_{j,\widetilde{j}}$ yields (in vector notation)

$$\mathbb{V}_{j,\widetilde{j}} = \text{Cov}\big(W'\mathbf{U}, \widetilde{W}'\widetilde{\mathbf{U}} \mid D\big) = W'\mathbb{E}[\boldsymbol{\varepsilon}\widetilde{\boldsymbol{\varepsilon}}']\,\widetilde{W}$$

with disturbances $\boldsymbol{\varepsilon}$ and $\widetilde{\boldsymbol{\varepsilon}}$ corresponding to domains j and $\widetilde{j}$, respectively. We estimate this quantity using the cluster-robust estimator

$$\widehat{\mathbb{V}}_{j,\widetilde{j}} := \sum_{h\in\mathcal{H}} \sum_{i\in h} \sum_{j\in h} \widetilde{w}_i \widehat{\widetilde{\varepsilon}}_j \widehat{\varepsilon}_j w_i,$$

which is consistent if $N^d/N \to c^d \in (0,1), d \in \{0,1\},\ H/N := |\mathcal{H}|/N \to c^H \in (0,1)$ as $N \to \infty$ and $|g| \leq c^G$ for all $h \in \mathcal{H},\ N \in \mathbb{N}$.[6]

Proof. For each estimator, the assumptions on N^d/N ensure that $(\widetilde{w}_i\widetilde{\varepsilon}_j H)$ is an $O_p(1)$ random variable, and so is $(w_i\varepsilon_j H)$, as $N \to \infty$. Because $|h|$ is bounded, $Z_h := \sum_{i\in h}\sum_{j\in h}(\widetilde{w}_i\widetilde{\varepsilon}_j H) \cdot (w_i\varepsilon_j H)$ is $O_p(1)$, too. Furthermore, $Z_h = o_p(1) + \sum_{i\in h}\sum_{j\in h}(\widetilde{w}_i\widehat{\widetilde{\varepsilon}}_j H)(w_i\widehat{\varepsilon}_j H)$, by consistency of the estimators under their respective assumptions. Hence,

$$\begin{aligned} H\widehat{\mathbb{V}}_{j,\widetilde{j}} &= HH^{-2}\sum_{h\in\mathcal{H}} Z_h + o_p(1) \\ &= o_p(1) + \mathbb{E}(Z_h). \end{aligned} \tag{A.9}$$

Likewise, we have that (again, cluster subscripts indicate observations corresponding to that cluster only)

$$\begin{aligned} W'\mathbb{E}[\boldsymbol{\varepsilon}\widetilde{\boldsymbol{\varepsilon}}']\,\widetilde{W} &= \mathbb{E}\left[\sum_{h\in\mathcal{H}} \widetilde{W}_h'\widetilde{\boldsymbol{\varepsilon}}_h\boldsymbol{\varepsilon}'\,W_h\right] \\ &= H\mathbb{E}\left[\widetilde{W}_h'\widetilde{\boldsymbol{\varepsilon}}_h\boldsymbol{\varepsilon}'\,W_h\right] \\ &= H\mathbb{E}\left[H^{-2}Z_h\right] \\ &= H^{-1}\mathbb{E}(Z_h). \end{aligned} \tag{A.10}$$

[6] w_i and ε_j or $\widetilde{w}_i$ and $\widetilde{\varepsilon}_j$ may also be vectors of fixed, finite dimension. Hence, the same cluster-robust estimator can be applied in the situation where observations cluster on both households h and ages $\mathcal{A}$.

□

Finally, we estimate the standard error of the aggregate treatment effect estimator (for estimator class c) as

$$\widehat{\sigma}_{c,\Sigma} := \sqrt{\sum_{j \in \mathcal{J}} \widehat{\mathbb{V}}_j + \sum_{j \in \mathcal{J}} \sum_{\widetilde{j} \in \mathcal{J} \setminus \mathcal{J}} \widehat{\mathbb{V}}_{j,\widetilde{j}}}.$$

A4. Alternative Evaluation Measures of Social Efficiency

We use the average net social benefit (NSB) as a measure of social efficiency. The NSB is based on basic economic principles. It compares the social benefit of a program to its social cost accounting for the welfare costs of taxation. We estimate the NSB for the average participant of PPP. Recently, Hendren and Sprung-Keyser (2020) and Finkelstein and Hendren (2020) advocate the use of the MVPF (marginal value of public funds) for evaluating and ranking social programs. The MVPF criterion compares a limited class of revenue-constant policies. It is based on accounting principles and favors programs that generate more government revenue per unit of expenditure. It ignores the scale of the programs and the marginal social cost of collecting revenue to finance them; it does not consider whether the societal utility possibility frontier is enhanced by programs.

García and Heckman (2022) provide a thorough discussion of measures of social efficiency. These authors note that, by definition, programs with the *same* net social benefit necessarily have MVPF $\geq$ BCR, but no special meaning can be attached to this algebraic consequence. A program with a high NSB can have a low MVPF and vice versa. We use NSB in this paper, as we are interested in whether the government budget should be expanded to fund programs like PPP and not whether adopting PPP holding total government revenue fixed is a good policy. We present the BCR as an intuitive measure of the social benefit per dollar of direct program cost, after verifying the inference on the social efficiency of the program using NSB.

A5. Details on the Monetization of the Life-Cycle Benefits and Costs

A5.1 Education

We estimate the present value generated by education costs. These costs include special education, K-12 education, and college.

K-12 and Special Education Costs

Records for K-12 education and special-education classes are almost entirely observed (118 out of 123 of PPP participants). We require no special analysis in terms of data preparation or additional estimation techniques. We assign an annual cost of 8,665 (2017 USD) per year of schooling and 18,803 (2017 USD) per year of special or remedial education year per participant. Estimates for expenses per student for regular education are taken from Grant and Lind (1978), corresponding to the school year 1975-1976. For that same period, Kakalik et al. (1981) provides a national ratio of current expenses per special education student to those per regular student of 2.17:1. We assume that this national ratio is comparable to the one applicable to Michigan. This factor is conservative because Kakalik et al. (1981) report estimates for the categories of special education that most likely apply to the PPP population (learning difficulties or different grades of mental retardation and emotional disturbance), which range between 2.3 and 3.8.

College Costs

We obtain estimates of college costs from Grant and Snyder (1986). We define the cost of college education for the PPP participants to society as the annual national expenditure of colleges per full-time equivalent student net of the average fees and in-state tuition of public colleges. We use estimates for the academic year 1982-1983, which is just after high school completion for most participants. We use tables 78 and 180 of Grant and Snyder

(1986)—13,768 (2017 USD) as the annual expenditure per college student in Michigan.

For the first two college enrollment periods, we have data on whether PPP participants were enrolled part-time. For every part-time enrollment, we only assign half of the annual cost. We do not distinguish between 2-year and 4-year college visits, even though expenditures per capita differ between these two institutions (see Grant and Snyder, 1986). We use the average between the expenditure of both types in our calculations.

Total Education Costs

We define the treatment effect on total costs of education to society simply as the sum of the two treatment effects described above.

A5.1.1 Estimation Specifics

Because of schooling timing differences between the control and treatment groups, we would need enrollment and matriculation records for each participant to monetize and appropriately discount college costs. However, college education data in the PPP sample was inconsistently recorded over different surveys. The age-27 and age-40 surveys have some inconsistencies. To minimize measurement error, we apply the data preparation algorithm outlined below. The preparation algorithm successfully resolves $\frac{2}{3}$ of the data inconsistencies.

We apply the **MD** and **OLS** estimators to the resulting complete case data. For **AIPW**, we supplement complete-case data with partially observed enrollment data. We proceed as follows. We consider college education costs, and deal with other educational costs analogously. For illustration purposes, we make the simplification that we first monetize and discount each individual path and then condition the discounted cost on covariates, instead of running age-wise regressions. Consider a partition of the population $\mathcal{P} = \mathcal{P}^c \cup \mathcal{P}^{ic} \cup \mathcal{P}^m$

Table A.4. Summary of Education Observations

Use Algorithm	Treatment Status $D = d$	Participants $\lvert\mathcal{P}\rvert$	Some College observed $\lvert\mathcal{P}^c \cup \mathcal{P}^{ic}\rvert$	College, partially obs. $\lvert\mathcal{P}^{ic}\rvert$	College data Missing $\lvert\mathcal{P}^m\rvert$
No	0	65	34	22	11
No	1	58	38	25	3
Yes	0	65	34	7	11
Yes	1	58	38	9	3

into individuals with complete educational records $\mathcal{P}^c$, those with partially observed educational records $\mathcal{P}^{ic}$, and those with fully missing records, $\mathcal{P}^m$. Note that for $i \in \mathcal{P}^{ic}$ we only know whether i was enrolled in college at some point, but cannot pinpoint when and for how long. Let $Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$ denote the total discounted cost of college education of person i, let R_i' be a binary indicator of $i \in \mathcal{P}^c \cup \mathcal{P}^{ic}$ (i.e., partially or fully observed Y_i) and let R_i be a binary indicator of $i \in \mathcal{P}^c$ (fully observed records). Similar to crime and earnings domains, Y_i is censored around 0, which we reflect with an enrollment or participation indicator, I_i. We can thus write $Y_i^d = I_i Y_i^{*,d}$, $Y_i^{*,d}$ as the total cost conditional on participation if the treatment status is fixed at d and I_i at 1. Let then $\ddot{Y}_i^d$ be a linear regression estimator for $\mathbb{E}[\, Y_i^{*,d} \mid Z_i, D_i = d, I_i = 1, R_i = R_i' = 1] = \mathbb{E}[\, Y_i^{*,d} \mid Z_i, D_i = d, I_i = 1]$. These are conditional (on Z_i) expected costs for those enrolled at some point ($I_i = 1$), estimated on the population $\{i : D_i = d, I_i = 1, R_i = R_i' = 1\}$, which is a subset of $\mathcal{P}^c$, hence observed. Second, let $\widetilde{Y}_i^d$ be a regression estimator of $\mathbb{P}(I_i = 1 \mid Z_i, D_i = d, R_i = R_i' = 1)\mathbb{E}[\, Y_i^{*,d} \mid Z_i, D_i = d, R_i = R_i' = 1] = \mathbb{P}(I_i = 1 \mid Z_i, D_i = d)\mathbb{E}[\, Y_i^{*,d} \mid Z_i, D_i = d]$.[7] Roughly speaking, $\ddot{Y}_i^d$ conditions on I_i and imputes outcomes for $i \in \mathcal{P}^{ic}$ while $\widetilde{Y}_i^d$ imputes outcomes for $i \in \mathcal{P}^m$. Then, set $\widehat{Y}_i^d := R_i' I_i \ddot{Y}_i^d + (1 - R_i') I_i \widetilde{Y}_i^d$. We use $\widehat{Y}_i^d$ in the **AIPW** estimator $\widehat{\Pi}_{\text{aipw}}$.

A5.1.2 Construction of Enrollment Profiles

To estimate the cost to society of providing PPP participants formal-education years such as college education, we need enrollment and matriculation records for each participant. However, the PPP sample's education data were inconsistently recorded in different surveys, especially the age-27 and age-40 survey, leading to disagreement in enrollment periods and considerable measurement error. To align data from different surveys, we clean it from obvious mistakes in data entry (e.g., by researching term lengths of colleges attended by

[7] We estimate the first factor using a Logit model. We estimate the second factor using a linear-regression model. Note that this procedure is slightly more formally articulated and otherwise akin to the one used in Section A5.2 for earnings.

individuals or removing participation in programs that do not fit the outcome). For all remaining issues, we use the following algorithm. Consider college enrollment again. Let $\theta_i = (t_i^{in}, t_i^{out}) \in \Theta_i,\ t_i^{in} < t_i^{out}$ be a pair of enrollment and matriculation dates for participant i and Θ_i be the set of all such pairs recorded in the PPP data for i. Now, a situation can arise in which $\theta_i \neq \theta'_i \in \Theta_i$ exist such that both periods overlap and hence conflict. If this is not the case, we say that θ_i, θ'_i are *consistent*. θ and θ' are consistent with one another, if and only if $(t_i^{out} \leq t_i^{in\prime}\ \text{ or }\ t_i^{out\prime} \leq t_i^{in})$. Furthermore, call a set $\widetilde{\Theta}_i \subseteq \Theta_i$ *consistent* if all its elements are pairwise consistent, and *maximally consistent*, if for any $\breve{\theta}_i \in \Theta_i \setminus \widetilde{\Theta}_i$, the set $\{\breve{\theta}_i\} \cup \Theta_i$ is not consistent. We assume that the true enrollment profile of i, $\Theta_i^* \subseteq \Theta_i$, is such a maximally consistent set. This equates to 1) assuming that there are no periods of parallel enrollments in two institutions and 2) that we trust all non-contradictory data. Third, we rely on longer periods of enrollment, so we do not accidentally discard, for instance, a college enrollment for a short job training. Define the overlap of two enrollment periods θ_i, θ'_i as

$$d(\theta_i, \theta'_i) = \frac{\min\left(\theta_i^{out} - \theta_i'^{\,in}, \theta_i'^{\,out} - \theta_i^{in}\right)}{\max\left(\theta_i^{out} - \theta_i^{in}, \theta_i'^{\,out} - \theta_i'^{\,in}\right)},$$

which is the share of the longer of the two periods that are overlapped by the shorter. We discard the shorter period whenever $d(\theta_i, \theta'_i) \geq 1-\gamma$ for some tuning parameter $\gamma \in (0,1)$. If we cannot discard the shorter period with this criterion, Θ_i cannot be made consistent and we interpret i's educational data as (partially) missing and set $R'_i = 1 - R_i = 0$. Likewise, if more than two periods overlap simultaneously or resolving conflicts in Θ_i in a different order yields different results, we set $R'_i = 1 - R_i = 0$. Algorithm 1 describes the exact procedure.

Algorithm 1: Preparing the education data for individual i with $R_i = 1$ (i.e., enrollment data was obtained.). Drop subscripts for brevity.

set $R' = 1$;
if Θ *is inconsistent (i.e., if there is a disagreement in education enrollment periods)* **then**
  define $\Theta' \subseteq \Theta$ as the subset of Θ containing all conflicting θs. pick some order of all $\theta_k \in \Theta'$. Let π^j be the jth permutation of the k's, $j = 1, ..., M$ with $M := |\Theta'|!$;
  set $j = 1$;
  while $j \leq M$ *or* $R' = 1$ **do**
    set $\Theta^j = \emptyset$, $\Theta'' = \Theta'$;
    take the ordering $\theta_{\pi^j(k)}, k = 1, ..., |\Theta'|$;
    while $\Theta'' \neq \emptyset$ **do**
      let $\breve{\theta}$ be the element with the lowest index in Θ'' according to ordering π^j;
      attempt to resolve the conflict of $\breve{\theta}$ using overlap criterion;
      if *conflict cannot be resolved or* $\breve{\theta}$ *overlaps with* ≥ 2 *elements* **then**
        set $R' = 0$, $\Theta = \emptyset$;
        exit;
      end
      remove $\breve{\theta}$ and elements discarded in the last step from Θ'', move $\breve{\theta}$ into Θ^j;
    end
    set $j = j + 1$;
  end
  if $\Theta^1 = ... = \Theta^M$ *and* $R' = 1$ **then**
    set $\Theta = \Theta^1$. This is i's final profile;
    exit.
  else
    set $\Theta = \emptyset$ and $R' = 0$;
    exit.
  end
end

A5.1.3 Education Costs for Intergenerational Impacts

To monetize costs of education for children, we use figures from National Center for Education Statistics and United States. Office of Educational Research and Improvement. Center for Education Statistics and Institute of Education Sciences (US) (2000), inflated to 2017 USD. For each student (full-time equivalent) and year, these are $29,181 for college, $11,376 for regular K-12 education and $24,685 for special education (assuming that the cost ratio for regular school to special education of 1:2.17 still holds).

A5.2 Labor Income

We obtain monthly observations on employment hours and wages directly from the PPP data. We examine the raw PPP data on employment histories and incarceration status to impute missing incomes.[8] We apply the **OLS** and **MD** estimators straight to the observed data as described. For the **AIPW** estimator, potential labor income is a censored variable only observed if the respective participant is employed. To construct **AIPW** first-stage estimates, we let $E_{i,a}$ be an indicator for whether individual i had some kind of employment at age a, then $Y_{i,a}$—which refers to annual labor income in this section—can be written as $Y_{i,a} = E_{i,a} Y^*_{i,a}$, where $Y^*_{i,a}$ is the potential labor earnings if i's employment status at age a were fixed at 1. Consequently, $\widehat{Y}^d_{i,a}$ is an estimate of $\mathbb{E}[Y^d_{i,a} \mid Z_i, D_i = d, R_{i,a} = 1, E_{i,a}]$ which equals $\mathbb{P}(E_{i,a} = 1 \mid Z_i, D_i = d, R_{i,a} = 1)\mathbb{E}[Y^{*,d}_{i,a} \mid Z_i, D_i = d, R_{i,a} = 1, E_{i,a} = 1]$ if unobserved and $E_{i,a}\mathbb{E}[Y^{*,d}_{i,a} \mid Z_i, D_i = d, R_{i,a} = 1, E_{i,a} = 1]$ if observed. We weight observations from adjoining ages using a normal kernel. Thus, our method forms imputations cross-sectionally and smoothes over the mean income path of the treatment and control groups.[9] Missing-data rates are fairly low, with 16% after data cleaning and preprocessing. Missing-data rates for treated and control participants are 14% and 18%, respectively.

For sensitivity analysis, we consider two additional methods for monetizing labor income. First, we interpolate the remaining missing values in the method described above. Second, we use the non-parametric matching method of García et al. (2020). We match each PPP participant with individuals in the National Longitudinal Study of the Young 1979 to impute missing values and forecast earnings after age 54, conditional on their observed earnings up

[8] For example, if observations of labor income are missing over a period during which a participant i was potentially incarcerated (we lack precise dates when individuals start serving their sentences, but we can bound the timing of prison sentences using their conviction year and length of sentence), we impute zero income during these periods. We do not assign a value to prison employment in our estimation since prison wages are negligible (less than $1 per hour).

[9] Note that we do not include income lags to predict $Y_{i,a}$. Preliminary calculations indicate that this does not add substance to our analyses. Especially, there is no indication that lagged observed income levels predict missing data after controlling for Z_i, D_i in a Logit model. Hence, our estimator assumption MAR II should be satisfied even without conditioning on income lags.

to that age. Appendix Table A.14 shows that our results are robust to using these two alternative methods.

A5.2.1 Labor Income Taxes

The increase in the earnings tax base due to higher labor income in the treatment group increases the state and federal tax base. We monetize these benefits by passing each labor-income observation through the appropriate tax function. To clarify this procedure, we replace age for time indices and write $Y_{i,t}$ as the individual- and time-specific labor income. We evaluate $Y_{i,t}$ in the appropriate tax function $g_{i,t}(Y_{i,t})$. The tax function outputs the amount of federal and state taxes (national average rate) that an individual would have to pay when earning labor income $Y_{i,t}$. We take the historical tax rates and deductibles at the state level from Citizens Research Council of Michigan (2021) and the historical average individual tax rates by income bracket at the federal level from Tax Policy Center (2020).

When choosing the function $g_{i,t}(\cdot)$ for each PPP participant, we consider their marital status. We proceed as follows:

1. If their marital status and spousal income is known, we form household income and calculate taxes based on it.

2. If their marital status is known and their spousal income is unknown but is observed in previous years, we assume that the spouse makes the last observed amount and proceed as in 1.

3. If their marital status is known and their spousal income is unknown and unobserved in previous years, we assume that the participant's labor income is half of the total household labor income and proceed as in 1. This may be inaccurate when computing

total household income privately. However, it is a good approximation for taxation purposes because labor-income taxation is (partly) based on household size.

4. If the individual is single or their marital status is missing, we simply assume individual taxation.

A5.2.2 Transfer Income

We calculate the benefit from transfers from the government to individuals as follows:

1. Individuals report whether they receive transfers from the government in the last two months from the following programs: temporary assistance to needy families, food stamps, child care subsidies, supplemental security income, unemployment insurance, general welfare assistance, disability payments, aid to families with dependent children, and any others. They report one figure per social program and we add up all amounts.

2. Individuals report for how many years in the last 15 years they have received money from any social programs in 1. We estimate that the average monthly amount they received in the last 15 years is equal to the amount received in 1. times the fraction of years in which they actually received transfers. This imputation carries the observed transfer payments backwards until the preceding interview, and accounts for welfare reforms that happened over the life-cycles of the PPP participants.

3. We discount and add up the amounts in 2. across each individual's life cycles and reverse the scale (i.e., multiply by -1). We reverse the scale because we consider a reduction in transfer income a benefit.

A5.3 Crime

We follow Hunt et al. (2017) and Hunt et al. (2019) and monetize the police and court costs per arrest by type of crime using estimates specific to Michigan. The police and court cost per arrest range from 2,371 (2017 US dollars) for drug-related crimes, driving offenses, vandalism, and similar crimes to 367,107 (2017 US dollars) for murder. We take correctional costs from United States Department of Justice (2010): 26,323 (2017 US dollars) per year spent in prison or jail, and 1,330 (2017 US dollars) per year of probation (due to monitoring). The prison cost is specific to Michigan, and the probation cost is national.

Victims of crime often sustain mental and physical healthcare costs, their property is damaged or stolen, become unable to work, or lose quality of life. To calculate the cost to the victims of crime, we proceed as follows. First, we follow standard practice in the criminology literature and inflate the number of arrests to address the disparity between the number of arrests and the number of crimes committed. Second, we monetize the material and quality-of-life losses to the crime victims. For the first step, we use several nationally representative datasets to estimate victimization-arrest inflation ratios. For the second step, we use the unit crime costs in Miller et al. (2020). Our estimates are conservative: under weak assumptions, possible measurement errors of victim costs are classical.[10] Appendix Table A.16 analyzes the sensitivity of our estimates to the several choices of monetization methodology in this section. We consider a scenario where the cost to crime victims is lower because stolen goods could be considered a transfer (without destruction of goods) between criminals and the victims.[11]

Let $\grave{Y}^{d}_{i,a}$ denote the total crime costs for individual i with treatment status fixed at d, at age a. The total crime costs include the sum of criminal justice system and victim costs flowing from all crimes committed at age a. Ideally, we would define $Y^{CJS\text{-}1,d}_{i,a,c}$ as the cost to the

[10]The strongest assumption is that, for a finely chosen crime typology and by treatment status, mean clearance rates are constant in criminal propensity and victim costs. We believe this assumption is justified since variation of underlying factors of clearance rates (e.g., police effort) is well captured by our detailed typology which differentiates by seriousness and nature of crime.

[11]The monetary cost to the victims includes medical and mental-health costs, loss of work, and lost and damaged property. These costs vary by crime type. Lost and damaged property is only a fraction of the total. We cannot separate "lost" from "damaged" in the sources that we use for monetizing crime.

criminal justice system that flow from each victimization (i.e., for investigating the crime, etc.) and the costs emanating from each arrest $Y_{i,a,c}^{CJS\text{-}2,d}$ (legal costs plus costs of keeping someone in prison if convicted) separately, where $c \in \mathcal{C}$ denotes a given crime category or crime type under consideration. Then, we would denote as $n_{i,a,c}^{d}$ the number of crimes committed, $\widetilde{n}_{i,a,c}^{d}$ the number of crimes with arrests for and $Y_{i,a,c}^{V,d}$ the average victim costs of the crimes committed by i at age a and in crime category $c \in \mathcal{C}$. The individual treatment effect on crime costs at age a in this scenario would be

$$\sum_{c \in \mathcal{C}} \left[n_{i,a,c}^{d} (Y_{i,a,c}^{V,d} + Y_{i,a,c}^{CJS\text{-}1,d}) + \widetilde{n}_{i,a,c}^{d} Y_{i,a,c}^{CJS\text{-}2,d} \right].$$

We simplify this ideal scenario and lump $Y_{i,a,c}^{CJS\text{-}1,d}$ and $Y_{i,a,c}^{CJS\text{-}2,d}$ into $Y_{i,a,c}^{CJS,d}$. We count criminal justice system costs by arrest and not by incidence. This attenuates our treatment-effect estimates, and leads to more conservative estimates. Our target individual-level parameter is

$$\grave{Y}_{i,a}^{1} - \grave{Y}_{i,a}^{0} := \tau_i = \underbrace{\left[\sum_{c \in \mathcal{C}} n_{i,a,c}^{1} Y_{i,a,c}^{V,1} - n_{i,a,c}^{0} Y_{i,a,c}^{V,0} \right]}_{:= \tau_i^{V}} + \underbrace{\left[\sum_{c \in \mathcal{C}} \widetilde{n}_{i,a,c}^{1} Y_{i,a,c}^{CJS,1} - \widetilde{n}_{i,a,c}^{0} Y_{i,a,c}^{CJS,0} \right]}_{:= \tau_i^{CJS}},$$

where τ_i^{V} and τ_i^{CJS} decompose the total treatment effect into victim and criminal justice system costs, and $\grave{Y}_{i,a}^{d} := \sum_{c \in \mathcal{C}} n_{i,a,c}^{d} Y_{i,a,c}^{V,d} + \widetilde{n}_{i,a,c}^{d} Y_{i,a,c}^{CJS,d}$ denotes what we define as the *true* cost of crime from individual i at age a (fixing $D_i = d$). Correspondingly, we set $\grave{Y}_{i,a}^{V,d} := \sum_{c \in \mathcal{C}} n_{i,a,c}^{d} Y_{i,a,c}^{V,d}$ and $\grave{Y}_{i,a}^{CJS,d} := \sum_{c \in \mathcal{C}} \widetilde{n}_{i,a,c}^{d} Y_{i,a,c}^{CJS,d}$.

For individual i fixed at treatment status $D_i = d$, we assign cost estimates for $\grave{Y}_{i,a,c}^{d}$ as follows. We do not observe the number of crimes of type c that i has committed ($n_{i,a,c}$). However, we observe the number of arrests ($\widetilde{n}_{i,a,c}$) that they have experienced at that age for that same crime. Therefore, we inflate the number of observed arrests by the ratio of total victimizations in the US to total number of arrests in the US (for that crime type), $\bar{\phi}_c$. We do not precisely observe the costs to the victims of i's committed crimes. Instead, we assign estimates of national mean victim costs, $\bar{y}_c^V$ for each crime category, c, as victim cost to the crime the PPP individual was arrested for. Criminal justice system costs on a state level (Michigan) are broken down and assigned to crime categories and summed with incarceration costs to form $\bar{y}_{i,a,c}^V$. This leads us to define

$$Y_{i,a}^d := \sum_c \left[\widetilde{n}_{i,a,c} \bar{\phi}_c \bar{y}_{i,a,c}^V + \widetilde{n}_{i,a,c} \bar{y}_c^{CJS} \right]$$

as the *observed* cost of crime flowing from i at age a. We describe our data and methods to construct $Y_{i,a}^d$ more precisely in Section A5.3.1.

We use inflation factors and cost estimates from data sources that match criminal activity in the PPP sample not only spatially but also temporally. We are not always able to estimate parameters that coincide with the structure and horizon of the PPP sample perfectly.

A5.3.1 Construction of Crime Costs

We define the cost of crime to society as the sum of costs to victims and costs to the criminal justice system (CJS costs). Victim costs include *medical* and *mental healthcare bills*, *damaged property*, *lost income* from employment disability, and *lost quality of life*. CJS costs include costs from *police investigation*, holding a *trial*, *incarceration*, and *probation*. We estimate both kinds of crime costs using the PPP crime data, supplemented with several national-level data files and cost estimates.

Crime Data Sources

The PPP crime data were collected from administrative, criminal records and reflect major checkpoints in a perpetrator's progress through the criminal justice system. For felonies, the data list every arrest, charge, conviction, prison sentence,[12] probation sentence and fine given to a PPP participant. They list every arrest, dropped charge, jail sentence, probation sentence, and fines for misdemeanors. For each arrest, charge, and conviction, it lists the set of crime types (e.g., non-negligent homicide, aggravated assault, motor vehicle theft) believed by the arresting officer or court to describe the crime event best.[13]

We supplement the PPP crime data with the *National Crime Victimization Survey* (NCVS), the *Uniform Crime Reports* (UCR), and the *National Judicial Reporting Program* (NJRP). We develop the common crime type categorization described below to harmonize each of these data sources with the PPP crime data. We present a short description of each data source in Table A.5.

In the PPP data, misdemeanors committed after the age 40 follow-up are classified using only four broad crime types: violent, property, drug-related, and other. To harmonize the data with the literature on crime costs to victims (which uses a finer set of crime types), we make the plausible assumption that all violent misdemeanors committed after age 40 are assaults, and all property misdemeanors committed after age 40 are larcenies. Tables A.7 and A.6 break down crime incidence for misdemeanors and felonies in control and treatment group, respectively.

Victim Costs

We sequentially tackle three challenges when estimating victim costs. First, only crimes that lead to arrests are observed. Hence, we inflate arrest counts and associated victim costs us-

[12]The data list the minimum and maximum prison sentence lengths assigned at conviction. We use the minimum sentence length as the actual time served, which is unknown.

[13]In total, there are 73 discrete crime types used in the data.

Table A.5. Description of Auxiliary Crime Data Sources

NCVS	The NCVS is a nationally representative, self-reported US survey on crime victimization at the household level. It provides detailed information on crimes, including those not reported to the police. We use the NCVS data to estimate total annual victimization in the US for six crime types: *rape/sexual assault*, *robbery*, *assault*, *burglary*, *larceny/theft*, and *motor-vehicle theft*. To minimize the burden on both surveyors and respondents, the NCVS allows surveyors to use one incident report to cover multiple incidents if they are similar in nature, occurred within a 6-month window, and are difficult for the respondent to distinguish. We include these "series crimes" but cap them at ten as suggested in Shook-Sa et al. (2015).
UCR	The UCR provides comprehensive arrest data for state and local agencies across the US beginning in 1980. It contains crimes to households, individuals, and businesses captured by most law enforcement agencies in the country. We use the UCR data to estimate total annual arrests in the US by crime type. The UCR includes all of the crime types surveyed in the NCVS plus "*murder*." We estimate the number of murder victimizations using the presumably more conservative "number of murders reported to the police" reported in the UCR. The UCR distinguishes *type-1 crimes* which are *murder, rape, assault, robbery, larceny, burglary,* and *motor-vehicle theft*, and type-2 crimes, those are sorted into 19 categories of less serious crimes.
NJRP	The NJRP collects detailed data on sentencing and offender characteristics from a nationally representative sample of convicted felons. We use the 2006 NJRP report (United States Department of Justice (2010)) to examine distributions of sentence types and lengths for each of the following crime types: *murder, rape, robbery, assault, burglary, larceny, fraud, motor vehicle theft, drug crimes,* and *other*. The 2006 NJRP is a reasonable choice since PPP participants' criminal activity was the highest during their late 20's/early 30's.

Table A.6. Administrative Data on Felonies, Summary Statistics

Crime Type	Arrests T	Arrests C	Convictions T	Convictions C	Prison Sentence T	Prison Sentence C	Years Incarcerated T	Years Incarcerated C	Probation Sentence T	Probation Sentence C
Murder	**0.02**	**0.05**	**0.02**	**0.03**	**0.02**	**0.03**	**7.50**	**32.50**	**0.00**	**0.00**
male	0.03	0.05	0.03	0.05	0.03	0.05	7.50	32.50	0.00	0.00
female	0.00	0.04	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Rape	**0.07**	**0.28**	**0.03**	**0.09**	**0.03**	**0.08**	**1.08**	**4.77**	**0.02**	**0.02**
male	0.12	0.46	0.06	0.15	0.06	0.13	1.08	4.77	0.03	0.03
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Robbery	**0.14**	**0.20**	**0.09**	**0.11**	**0.05**	**0.09**	**10.67**	**3.32**	**0.00**	**0.00**
male	0.24	0.31	0.15	0.15	0.09	0.13	10.67	3.18	0.00	0.00
female	0.00	0.04	0.00	0.04	0.00	0.04	.	4.00	0.00	0.00
Assault	**0.24**	**0.51**	**0.12**	**0.25**	**0.03**	**0.17**	**3.83**	**2.16**	**0.00**	**0.03**
male	0.42	0.85	0.21	0.41	0.06	0.28	3.83	2.16	0.00	0.05
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Burglary	**0.36**	**0.34**	**0.21**	**0.14**	**0.10**	**0.12**	**3.56**	**1.29**	**0.10**	**0.03**
male	0.64	0.54	0.36	0.23	0.18	0.21	3.56	1.29	0.18	0.05
female	0.00	0.04	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Larceny	**0.22**	**0.68**	**0.19**	**0.26**	**0.09**	**0.17**	**1.40**	**2.60**	**0.07**	**0.08**
male	0.39	1.08	0.33	0.38	0.15	0.26	1.40	2.85	0.12	0.10
female	0.00	0.08	0.00	0.08	0.00	0.04	.	0.08	0.00	0.04
Motor-Vehicle Theft	**0.03**	**0.08**	**0.02**	**0.03**	**0.00**	**0.02**	**.**	**0.25**	**0.00**	**0.02**
male	0.06	0.13	0.03	0.05	0.00	0.03	.	0.25	0.00	0.03
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Fraud	**0.47**	**0.11**	**0.07**	**0.05**	**0.02**	**0.03**	**1.00**	**0.75**	**0.05**	**0.02**
male	0.70	0.13	0.09	0.05	0.03	0.05	1.00	0.75	0.06	0.00
female	0.16	0.08	0.04	0.04	0.00	0.00	.	.	0.04	0.04
Vandalism	**0.02**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**.**	**.**	**0.00**	**0.00**
male	0.03	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Stolen Property	**0.07**	**0.09**	**0.05**	**0.11**	**0.05**	**0.08**	**1.19**	**1.93**	**0.02**	**0.02**
male	0.12	0.15	0.09	0.18	0.09	0.13	1.19	1.93	0.03	0.03
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Drug Offense	**0.36**	**0.54**	**0.28**	**0.31**	**0.16**	**0.20**	**6.14**	**2.27**	**0.05**	**0.11**
male	0.64	0.77	0.48	0.41	0.27	0.28	6.14	2.36	0.09	0.10
female	0.00	0.19	0.00	0.15	0.00	0.08	.	1.75	0.00	0.12
Disorderly Conduct	**0.10**	**0.26**	**0.14**	**0.11**	**0.05**	**0.06**	**2.06**	**1.63**	**0.02**	**0.00**
male	0.18	0.38	0.24	0.15	0.09	0.08	2.06	1.58	0.03	0.00
female	0.00	0.08	0.00	0.04	0.00	0.04	.	1.75	0.00	0.00
Miscellaneous	**0.26**	**0.15**	**0.07**	**0.08**	**0.03**	**0.00**	**1.58**	**.**	**0.02**	**0.02**
male	0.45	0.26	0.12	0.13	0.06	0.00	1.58	.	0.03	0.03
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Total	**2.36**	**3.28**	**1.28**	**1.55**	**0.64**	**1.05**	**3.97**	**3.22**	**0.34**	**0.32**
male	4.03	5.10	2.21	2.36	1.12	1.62	3.97	3.33	0.58	0.41
female	0.16	0.54	0.04	0.35	0.00	0.19	.	1.87	0.04	0.19

Note: This table summarizes the PPP administrative felony data. A column labeled with C displays the average in the control group. A column labeled with T displays the average in the treatment group. Prison and probation sentences are not assigned to individual citations at conviction, but to the bundle of (up to 7) citations associated with a given incident. In this table, we assign sentences to a specific crime type using the most serious crime type cited at conviction. Where possible, we use the hierarchy established by the UCR to determine crime seriousness (our ordering is given by the order of crime types in this table). The columns **Arrests**, **Convictions**, **Prison Sentence**, and **Probation Sentence** display average lifetime number of arrests, convictions, prison sentences, and probation sentences per participant. **Years Incarcerated** displays the average number of years incarcerated among participants who received a prison sentence.

Table A.7. Administrative Data on Misdemeanors, Summary Statistics

	Arrests		Convictions		Jail Sentence		Years Incarcerated		Probation Sentence	
Crime Type	T	C	T	C	T	C	T	C	T	C
Assault	**0.29**	**0.89**	**0.16**	**0.63**	**0.05**	**0.22**	**0.18**	**0.23**	**0.07**	**0.11**
male	0.45	1.15	0.24	0.82	0.09	0.28	0.18	0.28	0.09	0.18
female	0.08	0.50	0.04	0.35	0.00	0.12	.	0.08	0.04	0.00
Child Abuse	**0.02**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**.**	**.**	**0.00**	**0.00**
male	0.03	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Larceny	**0.40**	**0.66**	**0.31**	**0.62**	**0.14**	**0.26**	**0.41**	**0.12**	**0.03**	**0.03**
male	0.61	0.79	0.45	0.72	0.18	0.31	0.36	0.09	0.03	0.00
female	0.12	0.46	0.12	0.46	0.08	0.19	0.54	0.18	0.04	0.08
Burglary	**0.02**	**0.08**	**0.00**	**0.03**	**0.00**	**0.03**	**.**	**0.07**	**0.00**	**0.02**
male	0.03	0.13	0.00	0.05	0.00	0.05	.	0.07	0.00	0.03
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Fraud	**0.09**	**0.29**	**0.07**	**0.23**	**0.03**	**0.15**	**0.11**	**0.07**	**0.00**	**0.02**
male	0.06	0.36	0.06	0.28	0.06	0.18	0.11	0.08	0.00	0.03
female	0.12	0.19	0.08	0.15	0.00	0.12	.	0.06	0.00	0.00
Vandalism	**0.07**	**0.15**	**0.03**	**0.11**	**0.02**	**0.03**	**0.16**	**0.07**	**0.00**	**0.00**
male	0.09	0.21	0.03	0.13	0.03	0.05	0.16	0.07	0.00	0.00
female	0.04	0.08	0.04	0.08	0.00	0.00	.	.	0.00	0.00
Stolen Property	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**	**0.00**	**0.25**	**.**	**0.02**	**0.00**
male	0.03	0.03	0.03	0.03	0.03	0.00	0.25	.	0.03	0.00
female	0.00	0.00	0.00	0.00	0.00	0.00	.	.	0.00	0.00
Drug Offense	**0.22**	**0.42**	**0.16**	**0.37**	**0.05**	**0.15**	**0.13**	**0.26**	**0.05**	**0.02**
male	0.33	0.54	0.24	0.49	0.06	0.21	0.13	0.28	0.09	0.03
female	0.08	0.23	0.04	0.19	0.04	0.08	0.12	0.14	0.00	0.00
Disorderly Conduct	**0.59**	**0.72**	**0.57**	**0.65**	**0.10**	**0.20**	**0.11**	**0.14**	**0.05**	**0.11**
male	1.00	0.92	0.97	0.79	0.18	0.31	0.11	0.15	0.09	0.15
female	0.04	0.42	0.04	0.42	0.00	0.04	.	0.01	0.00	0.04
Driving Offense	**2.50**	**3.25**	**2.03**	**2.62**	**0.31**	**0.65**	**0.05**	**0.08**	**0.03**	**0.14**
male	3.18	4.21	2.64	3.38	0.42	0.95	0.05	0.08	0.00	0.15
female	1.60	1.81	1.24	1.46	0.16	0.19	0.05	0.13	0.08	0.12
Miscellaneous	**0.34**	**0.51**	**0.26**	**0.43**	**0.07**	**0.18**	**0.09**	**0.20**	**0.02**	**0.08**
male	0.48	0.69	0.36	0.62	0.12	0.26	0.09	0.24	0.03	0.13
female	0.16	0.23	0.12	0.15	0.00	0.08	.	0.02	0.00	0.00
Total	**4.55**	**6.98**	**3.60**	**5.69**	**0.79**	**1.88**	**0.15**	**0.14**	**0.28**	**0.51**
male	6.30	9.03	5.03	7.31	1.18	2.59	0.14	0.14	0.36	0.69
female	2.24	3.92	1.72	3.27	0.28	0.81	0.20	0.11	0.16	0.23

Note: This table summarizes the PPP administrative misdemeanor data. A column labeled with C displays the average in the control group. A column labeled with T displays the average in the treatment group. **Arrests**, **Convictions**, **Jail Sentence**, and **Probation Sentence** provide the average lifetime number of arrests, convictions, jail sentences, and probation sentences per participant; **Years Jailed** provides the average number of years jailed among participants who received a jail sentence.

ing national-level inflation factors. Second, we survey recent literature to locate estimates of mean crime costs per crime category. Third, we acknowledge that mean victimization costs per crime category might be very different in the PPP sample than in the general US population. Furthermore, mean victimization costs within crime categories might be different for treatment and control group (e.g., if control participants tended to commit more severe instances of crimes within any given category, we would underestimate the actual difference between the populations).

Crime Inflation Rates

We use several nationally representative datasets to construct victimization-arrest inflation factors to correct for unobserved PPP crimes. We estimate the total victim costs ($Y^V_{i,a}$) attributable to participant i at age a as the sum of crime-specific mean victim costs per arrest ($\widetilde{n}(i,a,c)\bar{y}^V_c$, where $\bar{y}^V_c$ is the mean victim cost of a crime in category c) inflated by the national victim-to-arrest ratio $\bar{\phi}_c$ ("VA-ratio" or victimization-inflation ratio) for that crime category (c) to account for unobserved crimes, that is $Y^V_{i,a} = \sum_{c \in \mathcal{C}} \widetilde{n}_{i,a,c}\bar{y}^V_c \cdot \bar{\phi}_c$.[14]

We use national average ratios of crime victimization to arrests for each of seven serious crime types reported in the NCVS: *murder, rape, robbery, assault, burglary, larceny,* and *motor-vehicle theft*. To impute other violent crimes, we calculate a "violent crime" victim-to-arrest ratio using the sum of rape, robbery, and assault. To impute other property crimes, we calculate a "property crime" victim-to-arrest ratio using the sum of burglary, larceny, and motor vehicle theft. Table A.9 displays our victim-arrest ratios estimated both on the data for the entire US and on data restricted to the Midwest.[15] We use the more conservative,

[14] We work with a time-invariant victimization-inflation ratio, drawing from crime data between 1995 and 2015. The actual victimization-inflation ratio has decreased over time. This inaccuracy leads to more conservative treatment-effect estimates. This is because 1) we will under-inflate crimes during peak times of criminal activity of the PPP sample, and 2) the control group maintains higher levels of criminal activity than the treatment group.

[15] The Midwest is formed by Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin.

smaller national estimates for our primary analyses.

The NCVS made significant methodological changes in 1993 (before which it was known as the *National Crime Survey*), making data from the years before 1993 incompatible with the years after. For this reason, and also because we do not know the exact year when PPP participants committed the crimes, we use victimization estimates averaged across all years in the NCVS (1994-2015) rather than estimating victimization every year. This approach is likely conservative as victimization appears to have been flat during the 1980s, trending downward throughout the 1990s, and then flattened out again in the 2000s (see figure A.3).

We calculate national arrest counts using the Uniform Crime Reports (UCR). Although available beginning in 1980, we restrict our attention to 1994 onward to match the NCVS availability. Additionally, as murder is unavailable in the NCVS, we use UCR data on clearances made by arrests to estimate the victim-arrest ratio for murder.[16]

Comparing victimization and arrest across datasets by crime type requires a common crime type categorization. We tabulate our harmonized crime categories across PPP data, UCR, and NCVS in Table A.10.

Unit Crime Costs

We draw average victim-cost estimates, $\bar{y}_c^V$, from Miller et al. (2020). Following the standard in the literature, Miller et al. (2020) uses a bottom-up approach to estimate victim costs per crime incident by type across several components. We calculate $\bar{y}_c^V$ as the sum of *medical*

[16]As participation in the UCR program is voluntary, many agencies do not submit complete arrest records for all 12 months. Following the methods suggested in the FBI's Crime in the US reports, Federal Bureau of Investigation (2020), we consider non-responding agencies (0-2 months reported) and partially responding agencies (3-11 months reported) separately. For non-responding agencies we estimate arrests using the arrest rate of agencies reporting 12 months in the same population size group. For agencies reporting 3-11 months of data, we simply inflate the total arrests by $12/N$ where N is the number of months reported.

costs, *mental* health costs, *work loss*, *property loss*, and *loss of quality of life*.[17] We exclude public services, adjudication and sanctioning, and perpetrator work loss.[18] Compared with prior studies (e.g., Cohen et al., 2004; McCollister et al., 2010), the estimates reported in Miller et al. (2020) are more conservative and cover a wider range of crime types. We provide an overview of average costs of crime to victims by crime type in Table A.11.

Costs to the Criminal Justice System

Costs to the criminal justice system are split between police, court, and correctional costs. We estimate the average police and court costs per arrest using an adjusted version of the methods developed in Hunt et al. (2017) and Hunt et al. (2019). Police and court costs vary by crime type. Specific to Michigan, we find a range between $2,371 (2017 USD) per arrest for UCR type-2 crimes (e.g., drug-related crimes, driving offenses, vandalism) and $367,107 (2017 USD) for murder.

Table A.8 shows the average maximum prison sentence assigned at conviction for both the PPP control and treatment groups. We compare these sentence lengths to the national average given in the National Judicial Reporting Program's 2006 report.

Hunt et al. (2019) develop a top-down approach to estimate the marginal cost of policing for UCR-type-1 crimes. Hunt et al. (2017) develop a similar approach to estimate the marginal cost to the court system for the same set of crimes. We adapt these approaches to estimate marginal costs to the police and court system in Michigan for each crime type used in the UCR (both type-1 and type-2). Correctional costs are added over time incarcerated/paroled

[17]To estimate the loss of quality of life, Miller et al. (2020) use a willingness-to-award approach based on jury verdicts and settlements. This approach leads to smaller estimates than the willingness-to-pay approaches used more frequently in the literature.

[18]Public services and adjudication and sanctioning costs are included as costs to the criminal justice system and our analyses on PPP earnings capture perpetrator work loss costs.

and vary by type of sentence. We take our annual costs of prison, jail, and probation from reports published by the Bureau of Justice Statistics (BJS).

Police and Court Costs

Hunt et al. (2019) use results from time-use surveys combined with information about a state's urban-rural composition and police-force role structure to allocate shares of annual law enforcement operating expenditure for various crime types. Similarly, Hunt et al. (2017) use results from time-use surveys and information on sentencing distributions to allocate shares of annual state expenditures on judicial and legal services to various crime types. We depart from their methods in two ways: 1) Because they estimate costs per crime reported to the police, they are limited to type-1 crimes for which the UCR records this information. We instead estimate costs per arrest, allowing us to expand our set of crimes to the complete set of type-1 and type-2 crimes; and 2) Secondly, they estimate costs for each state only for the year 2010. Instead, we use a time series of police expenditure information and arrests to estimate costs for each year between 1980 and 2015.[19]

We use the following formula (adjusted from Hunt et al. (2019)) to estimate $\omega_{c,t}^{police}$, the marginal cost of policing crime type c in year t in Michigan:

$$\begin{aligned} \omega_{c,t}^{police} =& \frac{\left(E_t \cdot \sum_r d_r\big(u \cdot p_{r,urban} + (1-u) \cdot p_{r,rural}\big)\right)}{A_{c,t}} \\ & \cdot \left(u \frac{A_{c,t}\tau_{c,urban}}{\sum\limits_{c'} A_{c',t}\tau_{c',urban}} + (1-u) \frac{A_{c,t}\tau_{c,rural}}{\sum_{c'} A_{c',t}\tau_{c',rural}} \right) \end{aligned} \tag{A.11}$$

where

- E_t is the annual law enforcement operating expenditure in Michigan in year t, taken from the Justice Expenditure and Employment Extracts (CJEE) for the years 1980-

[19] The PPP data cover 1973-2016. We use 1980 costs for years prior to 1980 and 2015 costs for the year 2016.

2015.[20] As recommended in Hunt et al. (2019), we inflate expenditures in the CJEE by 60% to account for the deadweight loss of taxation (20%) and the additional cost of providing employee benefits (40%).

- d_r, taken directly from Hunt et al. (2019), is the proportion of officers in Michigan assigned to role r, where officer roles include: general officers, community police officers, special task force officers, and detectives.
- u is the urban density of Michigan and is taken from the 1990, 2000, and 2010 US Census.[21]
- $p_{r,urban}$ and $p_{r,rural}$ are the proportions of time spent on crime by type r officers in urban and rural areas, respectively. We take the midpoint between minimum and maximum values given in Hunt et al. (2019).
- $\tau_{c,urban}$ and $\tau_{c,rural}$ are the number of hours spent on crime type c in urban and rural areas. We take the midpoint between minimum and maximum values given in Hunt et al. (2019).
- $A_{c,t}$ is the number of arrests of type c in Michigan in year t, calculated using the UCR arrest data for the years 1980-2015.

We estimate $\omega_{c,t}^{court}$, the marginal cost of policing crime type c in year t in Michigan using the following formula (adjusted from Hunt et al. (2017)):

$$\omega_{c,t}^{court} = E_t d \frac{p_c t_{c,fel} + (1 - p_c) t_{c,misd}}{\sum_{c'} (p_{c'} t_{c',fel} + (1 - p_{c'}) t_{c',misd}) A_{c,t}} \tag{A.12}$$

where

- E_t is the annual direct current judicial and legal expenditure in Michigan in year t, taken from the Justice Expenditure and Employment Extracts (CJEE) for the years 1982-2015.[22]
- d, taken directly from Hunt et al. (2017), is the proportion of cases in Michigan that are criminal cases
- p_c, taken directly from Hunt et al. (2017), is the proportion of type c crimes that are felonies. It is estimated as the proportion of type c crimes which result in a prison sentence. We set p_c equal to the felony proportion for larcenies (the lowest among type-1 crimes) for all UCR type-2 crimes.

[20]We use a linear interpolation to estimate missing years (1987, 1989, 1990, 1991, 2001, and 2003).

[21]The 1990 urban density is used for years prior to or including 1990, the 2000 urban density is used for years between 1990 and 2000, and the urban density for 2010 is used for years after 2000.

[22]We used linear interpolation to estimate missing years (1987, 1989, 1990, 1991, 2001, and 2003).

- $t_{c,fel}$ and $t_{c,misd}$ are the shares of criminal case time spent on felonies and misdemeanors of type c. We take the midpoint between minimum and maximum values given in Hunt et al. (2017) and then adjust such that the shares add to 1. We also set $t_{c,fel}$ equal to the share for felony larcenies (the type 1 crime with the lowest share) for all UCR-type-2 crimes.
- $A_{c,t}$ is the number of arrest of type c in Michigan in year t, calculated using the UCR arrest data for the years 1980-2015.

After adjusting for inflation costs, both police and court costs trend down through the 1980s and trend up beginning in 1990. Averages across all years are given in Table A.12.

Correctional Costs

We take our estimates of correctional costs from several reports published by the Bureau of Justice Statistics. US Department of Justice (1992) reports the annual cost of holding a person in prison in Michigan to be $31,222 (2017 USD), US Department of Justice (1984) and US Department of Justice (1990) report the annual cost of holding a person in jail in Michigan to be $27,064 (2017 USD) and $25,581 (2017 USD), respectively;[23] and US Department of Justice (1988) reports the annual cost of monitoring a person on probation nationally to be $1,330 (2017 USD).[24] We use the prison cost for all incarceration from felonies, the jail cost for incarceration from misdemeanors, and the probation cost for both felonies and misdemeanors.

[23] We use the average of these two estimates, $26,323 (2017 USD).

[24] We were unable to locate an estimate of probation costs specific to Michigan.

Table A.8. Average Prison Sentence Lengths for Felonies

Crime Type	**Treated, PPP**	**Control, PPP**	**National**
Murder	20.00	55.00	20.83
Rape	10.08	16.10	13.50
Robbery	33.33	15.10	8.42
Assault	4.50	3.83	5.17
Burglary	8.47	4.82	4.75
Larceny	2.54	7.72	3.17
Motor-Vehicle Theft	.	0.25	2.58
Fraud	14.00	1.75	3.75
Drug Offense	14.96	6.80	4.17

Note: This table reports the mean of maximum-sentence lengths assigned at conviction, which is available in both the PPP and NJRP data. National sentencing statistics are from the National Judicial Reporting Program (NJRP) for the year 2006.

Table A.9. Average Victimization-to-Arrest Ratios by Crime Type

Crime	**U.S.**	**Midwest**
Murder	1.65	2.35
Rape	7.48	9.13
Robbery	5.24	7.72
Assault	2.13	3.11
Burglary	12.83	18.78
Larceny	12.54	11.82
Motor Vehicle Theft	7.93	7.11
Violent Crime	2.97	4.36
Property Crime	12.19	12.32

Note: Violent crime includes rape, assault, and robbery. Property crime includes burglary, larceny, and motor vehicle theft.

Table A.10. Crime Categorization Across Data Sources

Category	PPP	UCRS	NCVS
Murder	Murder	Murder	
Rape	Rape	Forcible Rape	Completed Rape Attempted Rape
Robbery	Armed Robbery	Robbery	Robbery w/ Injury Robbery w/o Injury Attempted Robbery with Injury
Assault	Aggravated Assault Assault w/ Intent of Great Bodily Harm Assault w/ Intent to Murder Assault w/ Weapon Assault/Assault and Battery Aggravated Stalking Kidnapping	Aggravated Assault	Aggravated Assault with Injury Attempted Aggravated Assault
Burglary	Breaking and Entering Trespassing, Armed Home Invasion	Burglary	Burglary w/ Forcible Entry Burglary w/o Forcible Entry Attempted Forcible Entry
Larceny	Larceny (>$100) Larceny, in a Building Theft, of Rental Property Larceny (<$100) Larceny, from a Building Larceny, Shoplifting ($100)	Larceny	Purse Snatching Pocket Picking Theft Attempted Theft
Motor Vehicle Theft	Motor Vehicle Theft Unlawful Driving Away	Motor Vehicle Theft	Motor Vehicle Theft Attempted Motor Vehicle Theft

Table A.11. Average Costs of Crime to Victims by Crime Type

Crime Type	Medical	Mental Health	Lost Work	Lost Property	Quality of Life	Total
Assault	1,734	177	1,192	44	14,333	17,480
Child Abuse	9,708	3,891	1,443	7	43,415	58,464
Rape	1,835	4,108	4,575	176	152,683	163,377
Murder	12,735	11,976	1,828,638	197	5,150,836	7,004,382
Robbery	1,436	156	3,401	1,279	13,004	19,276
Fraud	0	0	57	1,854	0	1,911
Larceny/Theft	0	0	15	465	0	480
Burglary	0	0	23	1,641	0	1,664
Vandalism	0	0	0	390	0	390
Vehicle Theft	0	0	102	6,214	0	6,316

Note: All figures are taken from Miller et al. (2020) and inflated to 2017 USD.

Figure A.2. Victimization-to-Arrest Ratios, By Crime Type

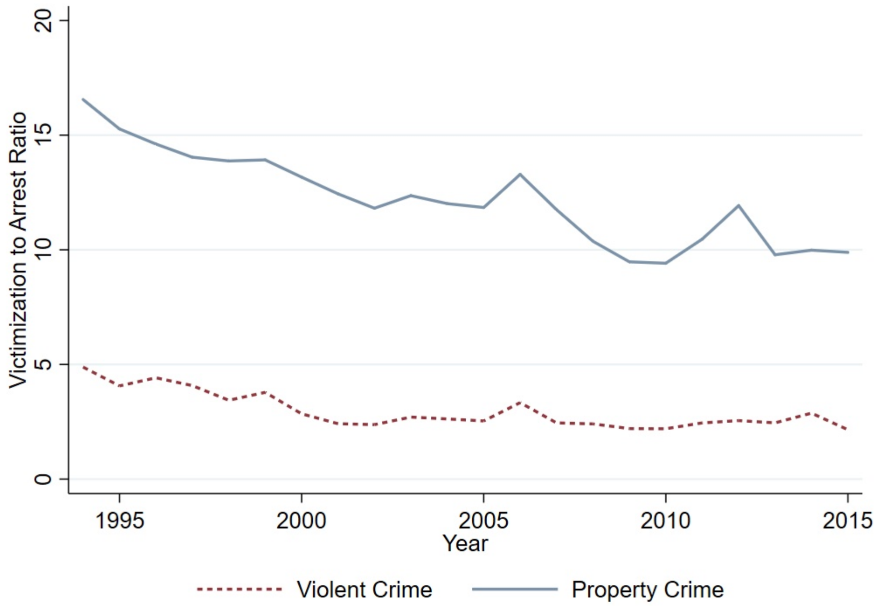
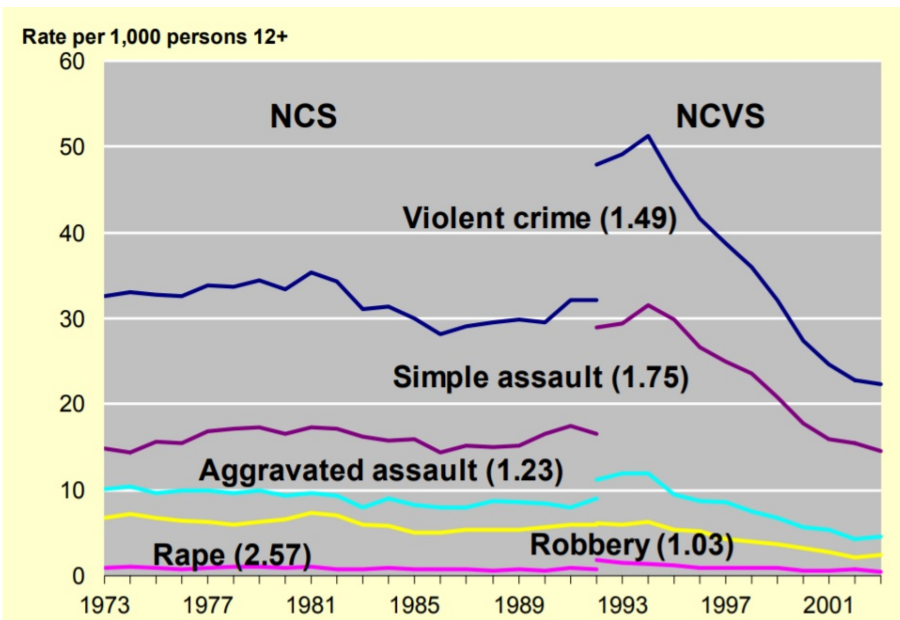


Table A.12. Cost of Crime to the Michigan Police and Court System, Averaged for the Period 1982-2015

Crime Type	**Police Cost**	**Court Cost**	**Combined Cost**
Murder	324,575	42,532	367,107
Rape	45,925	12,461	58,385
Robbery	6,907	2,603	9,510
Assault	26,666	2,015	28,681
Burglary	3,515	2,067	5,582
Larceny	3,162	1,772	4,933
Motor-Vehicle Theft	2,218	1,908	4,126
Type-2 Crimes	599	1,772	2,371

Note: All figures are inflated to 2017 USD.

Figure A.3. Violent Crime Rates, 1973-2003 and NCVS/NCS Ratio



Notes: The numbers in parentheses indicate the ratio between NCVS (post redesign) and NCS (pre-redesign) estimates of offense rates. Source: Rand (2006).

A5.4 Health

Overview. We simulate the health outcomes of participants in the age range between the age-27 and age-54 interviews using the Future Adult Model (FAM, Goldman, Ermini Leaf, and Tysinger, 2018). FAM simulates health-state transitions in the PSID.[25] We simulate the health outcomes after age 54 using the Future Elderly Model (FEM, Ermini Leaf et al., 2020; Goldman et al., 2018). FEM is analogous to FAM but is simulated in the HRS.[26] FAM does not simulate death because we observe this outcome in the age-54 interview. FEM simulates death as an outcome. If an individual dies, we assign all of their outcomes a monetary value of zero. We simulate the life-cycle trajectories of the PPP participants 1,000 times. The average across simulations is the point estimate of predicted health outcomes.

We compute governmental medical expenditures through programs like Medicaid and Medicare and private medical expenditures using the Medical Expenditure Panel Survey (MEPS) and the Medicare Current Beneficiary Survey (MCBS).[27] We do not predict health quality-adjusted life years (QALYs) before age 27 because the data required are not available in the PPP surveys. QALY gains generated by PPP are expected to be small when the participants are young. We predict QALYs after age 27 using the model in Goldman et al. (2018) adapted to the PSID for the outcomes of the FAM simulation. We proceed as in Goldman et al. (2018) for the outcomes of the FEM simulation. We use a yearly value of statistical life of 100,000 (2017 US dollars). QALYs adjust this yearly value to account for the burden of disease. Figures 1e and 1f display the cumulative average benefit from health over the life

[25] Panel Study of Income Dynamics (Institute for Social Research at the University of Michigan, 2019). We harmonize the PPP interviews at ages 27, 40, and 54 with the PSID. We use the age-27 interview as the initial period when starting the simulation.

[26] RAND Health and Retirement Study Version P public-use dataset and 2006-2014 biomarker sensitive datasets (RAND HRS Version P, 2016; University of Michigan and the National Institute on Aging, 2013, 2015a,b, 2016, 2017).

[27] We compute costs between age 27 and Medicare entry by applying MEPS-based models to the FAM and FEM predicted outcomes for the PPP participants. We assume that the participants enter Medicare at age 65 or after claiming disability for two consecutive years. After they enter Medicare, we follow an analogous procedure to compute their costs replacing MEPS with the 2007-2010 waves of MCBS.

cycle, adding up governmental and private medical expenditures and QALYs.[28]

Practical Details. For ages 18-27, we do not consider QALYs. Any resulting treatment-control differences would be minimal when individuals are that young. We predict medical costs from health outcomes reported by PPP participants in the age-19 and age-27 follow-ups harmonized with outcomes in MEPS. The only health outcome from the age-19 follow-up is the number of doctor visits due to of sickness in the past 12 months. From the age-27 follow-up, we use frequency of visiting a physician or other healthcare professional for routine physical examination, time since last visit to a physician or other healthcare professional for non-routine care, days spent in bed due to illness in the past 12 months, any hospital stays in the past 12 months, and treatment for any of these in the past five years: arthritis/joint disease, back injury, digestive tract disorder, fracture, hernia, eye disease, miscellaneous injury, pneumonia, psychosis, respiratory tract infection, and sprain.

We predict QALYs and medical costs between the age-27 and midlife follow-ups combining outcomes observed in the age-27, age-40, and midlife follow-ups, harmonizing outcomes across the PSID, MEPS, and MCBS data sets used in FAM. The PPP participant outcomes used for estimation in this age range are smoking history, body mass index, and diagnoses of cancer, diabetes, heart disease, hypertension, lung disease, and stroke.

After the midlife follow-up, we predict QALYs and medical costs over the remaining life course using outcomes observed at the age-40 and midlife follow-ups as the initial state variables for the FEM simulation. These outcomes are harmonized across the HRS, MEPS, and MCBS data sets. The PPP participant outcomes used for estimation are smoking history, body mass index, days spent in bed due to illness, physical activity level, and diagnoses of

[28] In addition to accounting for sampling variation as with all other monetized outcomes, our inference accounts for simulation error in the health predicted outcomes (i.e., forecasting error in the simulations of the life-cycle health trajectories). We provide details on this inference procedure in Appendix A3.

cancer, diabetes, heart disease, hypertension, lung disease, and stroke. These predictions also use physical measurements of the PPP participants' C-reactive protein, blood-sugar measures, cholesterol measures, and systolic blood pressure.

Sampling. We use the FAM and FEM models to monetize health outcomes. We use importance-weighted sampling to generate a distribution of outcomes simulated by FAM that is conditional on the outcomes observed in the PPP interviews at ages 40 and 54 (Ermini Leaf, 2017). We harmonize the PPP data up to the age-27 interview with MEPS data and predict costs using regression models estimated in the 2000-2010 MEPS waves.[29] We compute costs between age 27 and Medicare entry by applying MEPS-based models to the FAM and FEM predicted outcomes for the PPP participants. We assume that the participants enter Medicare at age 65 or after claiming disability for two consecutive years. After they enter Medicare, we follow an analogous procedure to compute their costs replacing MEPS with the 2007-2010 waves of MCBS.[30] We adjust the predicted values for average annual real medical cost growth using Congressional Budget Office (2007).

Estimation. These outcomes are governmental and private medical expenditure and quality-adjusted life years (QALYs). For QALYs, we assume that there is no treatment effect before age 30. We apply the **MD**, **OLS** and **AIPW** estimators as we do with the other outcomes. Missing data on monetized health outcomes is rare (3.2% in the pooled sample of all individuals of all ages). The instances of missing data are generated by missing values in the FAM or FEM models' inputs. These models are simulated to forecast the life-cycle trajectories of several health outcomes, to then monetize them into the expenditures and QALYs. We average by age and individual over the 1,000 simulated life-cycle trajectories, conditional on all data available by age 54. Each simulated life-cycle path ends with participants' simulated

[29] Agency for Healthcare Research and Quality (2003, 2004a,b, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014).

[30] Centers for Medicare and Medicaid Services, U.S. Department of Health (2011a,b, 2012, 2013).

death at a random age, which we truncate at age 99. We treat observed deaths before age 54 as a conditioning variable. Therefore, expected medical costs and QALYs of participants who die before age 54 are set to 0.

Additional Cost Details. Medicaid costs are shared between states and the federal government. The federal share for the states is determined by each state's Federal Medical Assistance Percentage (FMAP; KFF, 2012). MEPS and MCBS provide data on Medicaid expenditures—state expenditures excluding Medicaid and federal expenditures excluding Medicaid. We estimate costs and then allocate the Medicaid amount to the state and federal amounts using Michigan's FMAP. Historical FMAP values for 1976-2004 were published in the Federal Register.[31] Historical and estimated values for 2005-2022 were obtained from KFF's State Health Facts database.[32] After 2022, we assume that Michigan's FMAP remains constant at 65%. This percentage is the four-year average before the Covid-19 pandemic started in 2020. Using the FMAP provides us with a conservative estimate of the federal share of Medicaid expenditures because some of the PPP participants might qualify for an enhanced FMAP at various times during their lifetime. We do not track eligibility for these enhanced FMAPs.

For health, which we monetize up to realized or forecasted death, the forecasting model that we use considers public medical costs as the only post-retirement transfer from the government to individuals. Unfortunately, our data regarding retirement and savings accounts is limited to further analyzing pensions or social security payments.

[31] DHEW Federal Financial Participation in State Assistance Expenditures (1979); DHEW State Assistance Expenditures (1974, 1976) and DHHS Federal Financial Participation in State Assistance Expenditures (1980, 1982, 1984, 1986, 1987a,b, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1996, 1997a,b, 1999, 2000a,b, 2001, 2002).

[32] KFF's State Health Facts (2021) sourced from DHHS Adjusted Federal Medical Assistance Percentage (FMAP) Rates (2011); DHHS Federal Financial Participation in State Assistance Expenditures (2003a,b, 2004, 2005, 2006, 2010, 2011, 2012, 2014a,b, 2015, 2016); DHHS Federal Matching Shares for Medicaid (2017, 2018, 2019, 2020); DHHS Implementation of Section 5001 of the American Recovery and Reinvestment Act of 2009 (2009, 2010).

Other Qualifications. Changing preferences due to improved health is another potential benefit source we do not consider here. We do not have data to quantify such benefit. The nature of the referred preference changes is not established in the literature. For instance, if poorer health reduces the marginal utility of consumption, individuals could decide to spend more in out-of-pocket medical expenses. This implies that private insurance premiums and taxation to fund public medical expenditures are higher than optimal. Accounting for these inefficiencies could attenuate our estimate of the health present value. Finkelstein et al. (2009) review the challenges of estimating health effects on the marginal utility of consumption. Finkelstein et al. (2013) and Blundell et al. (2020) are related studies.

A6. Sensitivity Analysis and Additional Estimates

In this appendix, we present additional estimates and sensitivity analysis. First, we present estimates that include the gain in terms of QALYs from lower crime victimization in Appendix Table A.13. Second, we present sensitivity analysis of estimation choices using Table 4 as a benchmark (these estimates are in Appendix Table A.14). Third, we present inference on our preferred benefit-cost ratio estimates in Table 4 using alternative methods (this exercise is in Appendix Table A.15). We also present sensitivity analysis to the use of alternative estimators. Finally, we present sensitivity analysis to the use of externally supplied parameters to compute our preferred benefit-cost ratio in Table 4 (this exercise is in Appendix Table A.16).

Table A.13. Benefit-Cost Ratio for the Participants of PPP, Main and Additional Results

Present Values in 1,000s of 2017 USD	**Baseline in Table 4** **Estimate**	**Add Crime QALY Costs** **Estimate**
Benefit-Cost Ratio		
Baseline Program Cost	***8.98***	***17.23***
Subtract Deadweight Loss (50%)	***5.98***	***11.48***

Note: This table summarizes our preferred estimate of the benefit-cost ratio—as defined in Equation (5). It also presents additional estimates that include the gain in terms of QALYs from lower crime victimization. We bold (italicize) estimates when they are significant at the 10% (5%) based on bias-corrected accelerated bootstrap p-values. The null hypothesis for the benefit-cost ratio is that it is less than or equal to 1. The estimates rely on the OLS estimator explained in Section 3, which adjusts for compromises in the randomization protocol, attrition, and item non-response. The estimates are discounted to the year in which the program started using a rate of 3%. We show benefit-cost ratios using the baseline program cost (21,151 of 2017 US dollars), as well as using the baseline program cost multiplied by 1.5 to account for the deadweight that would be generated by collecting the taxes required to fund the program.

Table A.14. Life-Cycle Present Value and Benefit-Cost Ratio for the Participants of PPP, Sensitivity Analysis of Estimation Choices

Present Values in 1,000s of 2017 USD	**Income Model**		**Life-Cycle Segment**				**Set Outcome 0**			
(Baseline)	*(Observed,* 16 − 40*)*		*(Full records for Education,* 16 − 40 *for Income and Crime, 30-Death for Health)*				*(None)*			
	(16 − 40)	(16 − 60)								
Change from Baseline	*Linear Interpolation*	*García et al. (2020)*	0 − 20	21 − 40	41 − 54	55*–Death*	*Education*	*Income*	*Crime*	*Health*
Education Present Value										
Total	0.27	0.27	−1.67	2.40				0.27	0.27	0.27
[s.e.]	[2.42]	[2.42]	[2.08]	[1.59]				[2.42]	[2.42]	[2.42]
Labor Income Present Value										
Total	66.99	62.97	5.65	56.44	−5.86	−0.57	61.58		61.58	61.58
[s.e.]	[29.20]	[40.80]	[6.10]	[28.60]	[18.91]	[1.50]	[30.06]		[30.06]	[30.06]
Crime Present Value										
Criminal Justice System Cost	19.21	19.21	−9.68	28.90	−0.81		19.21	19.21		19.21
Monetary Cost to Victims	60.13	60.13	−28.81	88.96	−1.72		60.13	60.13		60.13
Total	79.34	79.34	−38.49	117.86	−2.53		79.34	79.34		79.34
[s.e.]	[60.30]	[60.30]	[27.68]	[50.79]	[2.67]		[60.30]	[60.30]		[60.30]
Health Present Value										
Government Expenditure	−2.00	−2.00	0.02	2.29	0.46	−5.02	−2.00	−2.00	−2.00	
Private Expenditure	−9.00	−9.00	0.02	−2.34	−2.60	−4.10	−9.00	−9.00	−9.00	
Quality-Adjusted Life Years	59.69	59.69	0.00	24.49	18.32	21.66	59.66	59.66	59.66	
Total	48.69	48.69	0.04	24.45	16.18	12.54	48.69	48.69	48.69	
[s.e.]	[72.61]	[72.61]	[0.21]	[26.49]	[26.33]	[25.44]	[72.61]	[72.61]	[72.61]	
Total Present Value	195.29	191.27	−34.48	201.15	7.80	11.97	189.60	128.30	110.54	141.19
[s.e.]	[99.36]	[109.92]	[29.55]	[65.47]	[37.70]	[25.90]	[100.76]	[87.90]	[84.19]	[71.76]
Benefit-Cost Ratio										
Baseline Program Cost	9.23	9.04	−1.63	9.51	0.37	0.57	8.96	6.07	5.23	6.68
[s.e.]	[4.70]	[5.20]	[1.40]	[3.10]	[1.78]	[1.22]	[4.76]	[4.16]	[3.98]	[3.39]
Subtract Deadweight Loss (50%)	6.16	6.03	−1.09	6.34	0.25	0.38	5.98	4.04	3.48	4.45
[s.e.]	[3.13]	[3.46]	[0.93]	[2.06]	[1.19]	[0.82]	[3.18]	[2.77]	[2.65]	[2.26]

Note: The columns in this table summarize specifications that vary one aspect of our preferred specification, Table 4. We vary the strategy for calculating the labor-income benefits—from the baseline using observation only to interpolating as explained in Section 3 or interpolating and extrapolating using the method in García et al. (2020) and the age-range considered. We also consider specifications setting the present value to 0 for each of the outcomes, one at a time. Empty entries indicate that component is set to 0 in column specification. The standard errors in brackets are bootstrapped and clustered at the household level.

Table A.15. Life-Cycle Present Value and Benefit-Cost Ratio, Additional Robustness Checks for Inference and Estimators

	Main Specification		**Alternative *p*-values**			
	Estimate	*p*-value	Analytic	Bootstrap		
Present Values in 1,000s of 2017 USD				Simple	Trimmed	Studentized
Panel (a): MD						
Benefit-Cost Ratios						
Baseline Program Cost	***8.03***	.043	.066	.077	.062	.089
Subtract Deadweight Loss (50%)	**5.35**	.057	.081	.093	.077	.100
Panel (b): OLS (Table 4)						
Benefit-Cost Ratios						
Baseline Program Cost	***8.98***	.024	.039	.060	.046	.066
Subtract Deadweight Loss (50%)	***5.98***	.034	.049	.072	.058	.075
Panel (c): AIPW						
Benefit-Cost Ratios						
Baseline Program Cost	**8.72**	.063	.061	.151	.076	.085
Subtract Deadweight Loss (50%)	**5.81**	.081	.074	.167	.090	.093

Note: This table presents benefit-cost ratio estimates and the p-values corresponding to our main specification (Table 4). We also provide estimates for the two alternative estimators discussed in Section 3. The baseline p-values are bias-corrected accelerated and bootstrapped. Based on these p-values, we bold (italicize) the point estimates when they are significant at the 10% (5%). The null hypothesis for the benefit-cost ratio is that it is less than or equal to 1. The alternative p-values are analytic asymptotic or based on the different bootstrap procedures indicated in the label. Details on all of the inference procedures are in Appendix A3.

Table A.16. Benefit-Cost Ratio, Robustness Checks for Externally Supplied Parameters

		Pessimistic Scenarios		**Optimistic Scenarios**	
Baseline Benefit-Cost (BC) Ratio: ***5.98*** Confidence Interval: [1.37 , 12.91]	Baseline Parameter (1)	Parameter Multiplier (2)	Reestimated BC-Ratio (3)	Parameter Multiplier (4)	Reestimated BC-Ratio (5)
Education					
K-12			***5.90***		***6.07***
	$8,665	×1.5	[1.27 , 12.79]	×0.5	[1.44 , 13.00]
Special Education			***5.93***		***6.04***
	$18,803	×0.5	[1.30 , 12.80]	×1.5	[1.40 , 13.00]
College			***6.02***		***5.95***
	$13,768	×1.5	[1.40 , 12.93]	×0.5	[1.33 , 12.89]
Crime					
Incarceration	Listed in Appendix A5.3		***5.87***		***6.10***
		×0.5	[1.26 , 12.44]	×1.5	[1.69 , 13.24]
Cost to Victim			***5.04***		***6.93***
		×0.5	[0.26 , 10.30]	×1.5	[1.90 , 15.79]
Criminal Justice			***5.80***		***6.17***
		×0.5	[1.21 , 12.68]	×1.5	[1.52 , 13.12]
Victimization Inflation			***5.04***		***6.93***
		×0.5	[0.26 , 10.30]	×1.5	[1.90 , 15.79]
Health					
Value of Statistical Life			***5.04***		***6.92***
	$100,000	×0.5	[1.68 , 10.98]	×1.5	[0.61 , 15.92]
Cash-Flow Discount Rate			***3.18***		***17.15***
	3%	$\times\frac{5}{3}$	[0.61 , 6.80]	×0	[1.64 , 40.11]

Note: Column [1] presents the externally supplied parameters that we use for monetizing treatment effects in Table 4. These parameters are *per annum* costs or benefits. We then present pessimistic (optimistic) scenarios where the parameters are multiplied by 1.5 (0.5) if they represent a cost and multiplied by 0.5 (1.5) if they represent a benefit. We reestimate our baseline estimate of the benefit-cost ratio when perturbing individually each parameter (one at a time). We display two-sided bias-corrected accelerated bootstrap 90% confidence intervals in brackets, clustered at the household level. We bold (italicize) benefit-cost ratios when they are significant at the 10% (5%) based on bias-corrected accelerated bootstrap p-values. The null hypothesis for the benefit-cost ratio is that it is less than or equal to 1.

Appendix References

Agency for Healthcare Research and Quality (2003). MEPS HC-050: 2000 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2004a). MEPS HC-060: 2001 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2004b). MEPS HC-070: 2002 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2005). MEPS HC-079: 2003 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2006). MEPS HC-089: 2004 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2007). MEPS HC-097: 2005 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2008). MEPS HC-105: 2006 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2009). MEPS HC-113: 2007 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2010). MEPS HC-121: 2008 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2011). MEPS HC-129: 2009 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2012). MEPS HC-138: 2010 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2013). MEPS HC-147: 2011 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Agency for Healthcare Research and Quality (2014). MEPS HC-155: 2012 Full Year Consolidated Data File. Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality.

Bell, R. M. and D. F. McCaffrey (2002). Bias Reduction in Standard Errors for Linear Regression With Multi-Stage Samples. *Survey Methodology 28*(2), 169–182.

Blundell, R., M. Borella, J. Commault, and M. De Nardi (2020). Old Age Risks, Consumption, and Insurance. NBER Working Paper w27348, National Bureau of Economic Research.

Centers for Medicare and Medicaid Services, U.S. Department of Health (2011a). Medicare Current Beneficiary Survey, 2007 Cost and Use Data. U.S. Department of Health and Human Services.

Centers for Medicare and Medicaid Services, U.S. Department of Health (2011b). Medicare Current Beneficiary Survey, 2008 Cost and Use Data. U.S. Department of Health and Human Services.

Centers for Medicare and Medicaid Services, U.S. Department of Health (2012). Medicare Current Beneficiary Survey, 2009 Cost and Use Data. U.S. Department of Health and Human Services.

Centers for Medicare and Medicaid Services, U.S. Department of Health (2013). Medicare Current Beneficiary Survey, 2010 Cost and Use Data. U.S. Department of Health and Human Services.

Citizens Research Council of Michigan (2021). Michigan Historic Personal Income Tax Revenue: 1968 to Present. `https://crcmich.org/almanac/historic-personal-income-tax`. [Online; accessed 14-January-2021].

Cohen, M. A., R. T. Rust, S. Steen, and S. T. Tidd (2004). Willingness-to-Pay for Crime Control Programs. *Criminology 42*(1), 89–110.

DHEW Federal Financial Participation in State Assistance Expenditures (1979). DHEW Federal Financial Participation in State Assistance Expenditures, 44 Fed. Reg. 10553.

DHEW State Assistance Expenditures (1974). DHEW State Assistance Expenditures, 39 Fed. Reg. 33020.

DHEW State Assistance Expenditures (1976). DHEW State Assistance Expenditures, 41 Fed. Reg. 44879.

DHHS Adjusted Federal Medical Assistance Percentage (FMAP) Rates (2011). DHHS Adjusted Federal Medical Assistance Percentage (FMAP) Rates for the Second and Third Quarters of Fiscal Year 2011 (FY11), 76 Fed. Reg. 32204.

DHHS Federal Financial Participation in State Assistance Expenditures (1980). DHHS Federal Financial Participation in State Assistance Expenditures, 45 Fed. Reg. 79582.

DHHS Federal Financial Participation in State Assistance Expenditures (1982). DHHS Federal Financial Participation in State Assistance Expenditures, 47 Fed. Reg. 56401.

DHHS Federal Financial Participation in State Assistance Expenditures (1984). Federal Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1985 Through September 30, 1987, 49 Fed. Reg. 46957.

DHHS Federal Financial Participation in State Assistance Expenditures (1986). Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or, Disabled Persons, 51 Fed. Reg. 39915.

DHHS Federal Financial Participation in State Assistance Expenditures (1987a). Federal Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1988 Through September 30, 1989, 52 Fed. Reg. 41506.

DHHS Federal Financial Participation in State Assistance Expenditures (1987b). Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1987 Through September 30, 1988, 52 Fed. Reg. 12253.

DHHS Federal Financial Participation in State Assistance Expenditures (1988). Federal Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1989 Through September,30, 1990, 53 Fed. Reg. 43477.

DHHS Federal Financial Participation in State Assistance Expenditures (1989). Federal Matching Shares for Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1990 Through September 30, 1991, 54 Fed. Reg. 49358.

DHHS Federal Financial Participation in State Assistance Expenditures (1990). Federal Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1991 through September 30, 1992, 55 Fed. Reg. 48170.

DHHS Federal Financial Participation in State Assistance Expenditures (1991). 56 Fed. Reg. 58249.

DHHS Federal Financial Participation in State Assistance Expenditures (1992). Federal Matching Shares for Aid to Families With Dependent Children, Foster Care and Adoption Assistance, Job Opportunities and Basic Skills Training, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1993, Through September 30,1994, 57 Fed. Reg. 57837.

DHHS Federal Financial Participation in State Assistance Expenditures (1993). Federal Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1994 Through September 30, 1995, 58 Fed. Reg. 66362.

DHHS Federal Financial Participation in State Assistance Expenditures (1994). Federal Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1995 through September 30, 1996, 59 Fed. Reg. 39407.

DHHS Federal Financial Participation in State Assistance Expenditures (1996). Federal Matching Shares for Aid to Families With Dependent Children, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1996 Through September 30, 1997, 61 Fed. Reg. 8288.

DHHS Federal Financial Participation in State Assistance Expenditures (1997a). Federal Matching Shares for Temporary Assistance to Needy Families, Medicaid, Aid to Needy Aged, Blind, or Disabled Persons and for the New Children's Health Insurance Programs for October 1, 1998 Through September 30, 1999, 62 Fed. Reg. 62613.

DHHS Federal Financial Participation in State Assistance Expenditures (1997b). Federal Matching Shares for Temporary Assistance to Needy Families, Medicaid, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 1997 Through September 30, 1998, 62 Fed. Reg. 4293.

DHHS Federal Financial Participation in State Assistance Expenditures (1999). Federal Matching Shares for State Children's Health Insurance Programs and for Selected Portions of State Medicaid Programs for October 1, 1999 Through September 30, 2000; Correction, 64 Fed. Reg. 52095.

DHHS Federal Financial Participation in State Assistance Expenditures (2000a). Federal Matching Shares for Temporary Assistance to Needy Families, Medicaid, Aid to Needy Aged, Blind, or Disabled Persons and for the State Children's Health Insurance Program for October 1, 2000 through September 30, 2001, 65 Fed. Reg. 8979.

DHHS Federal Financial Participation in State Assistance Expenditures (2000b). Federal Matching Shares for Temporary Assistance to Needy Families, Medicaid, Aid to Needy Aged, Blind, or Disabled Persons and State Children's Health Insurance Program for October 1, 2001 through September 30, 2002, 65 Fed. Reg. 69560.

DHHS Federal Financial Participation in State Assistance Expenditures (2001). Federal Matching Shares for Medicaid, the State Children's Health Insurance Program, and Aid to Needy, Aged, Blind, or Disabled Persons for October 1, 2002 Through September 30, 2003, 66 Fed. Reg. 59790.

DHHS Federal Financial Participation in State Assistance Expenditures (2002). Federal Matching Shares for Medicaid, the State Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2003 Through September 30, 2004, 67 Fed. Reg. 69223.

DHHS Federal Financial Participation in State Assistance Expenditures (2003a). Federal Matching Shares for Medicaid, the State Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2004 Through September 30, 2005, 68 Fed. Reg. 67676.

DHHS Federal Financial Participation in State Assistance Expenditures (2003b). Temporary Increase of Federal Matching Shares for Medicaid for the Last 2 Calendar Quarters of Fiscal Year 2003 and the First 3 Quarters of Fiscal Year 2004, 68 Fed. Reg. 35889.

DHHS Federal Financial Participation in State Assistance Expenditures (2004). Federal Matching Shares for Medicaid, the State Children's Health Insurance Program, and Aid To Needy Aged, Blind, or Disabled Persons for October 1, 2005 Through September 30, 2006, 69 Fed. Reg. 68370.

DHHS Federal Financial Participation in State Assistance Expenditures (2005). Federal Matching Shares for Medicaid, the State Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2006 Through September 30, 2007, 70 Fed. Reg. 71856.

DHHS Federal Financial Participation in State Assistance Expenditures (2006). Federal Matching Shares for Medicaid, the State Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2007 Through September 30, 2008, 71 Fed. Reg. 66209.

DHHS Federal Financial Participation in State Assistance Expenditures (2010). DHHS Adjusted Federal Medical Assistance Percentage (FMAP) Rates for the Second and Third Quarters of Fiscal Year 2011 (FY11), 76 Fed. Reg. 32204.

DHHS Federal Financial Participation in State Assistance Expenditures (2011). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2012 Through September 30, 2013, 76 Fed. Reg. 74061.

DHHS Federal Financial Participation in State Assistance Expenditures (2012). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2013 Through September 30, 2014, 77 Fed. Reg. 71420.

DHHS Federal Financial Participation in State Assistance Expenditures (2014a). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2014 Through September 30, 2015, 79 Fed. Reg. 3385.

DHHS Federal Financial Participation in State Assistance Expenditures (2014b). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2015 Through September 30, 2016, 79 Fed. Reg. 71426.

DHHS Federal Financial Participation in State Assistance Expenditures (2015). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2016 Through September 30, 2017, 80 Fed. Reg. 73779.

DHHS Federal Financial Participation in State Assistance Expenditures (2016). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2017 Through September 30, 2018, 81 Fed. Reg. 80078.

DHHS Federal Matching Shares for Medicaid (2017). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2018 Through September 30, 2019, 82 Fed. Reg. 55383.

DHHS Federal Matching Shares for Medicaid (2018). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2019 Through September 30, 2020, 83 Fed. Reg. 61157.

DHHS Federal Matching Shares for Medicaid (2019). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2020 Through September 30, 2021, 84 Fed. Reg. 66204.

DHHS Federal Matching Shares for Medicaid (2020). Federal Matching Shares for Medicaid, the Children's Health Insurance Program, and Aid to Needy Aged, Blind, or Disabled Persons for October 1, 2021 Through September 30, 2022, 85 Fed. Reg. 76586.

DHHS Implementation of Section 5001 of the American Recovery and Reinvestment Act of 2009 (2009). Adjustments to the Third and Fourth Quarters of Fiscal Year 2009 Federal Medical Assistance Percentage Rates for Federal Matching Shares for Medicaid and Title IV–E Foster Care, Adoption Assistance and Guardianship Assistance Programs, 74 Fed. Reg. 64697.

DHHS Implementation of Section 5001 of the American Recovery and Reinvestment Act of 2009 (2010). Adjustments to the Fourth Quarter of Fiscal Year 2010 Federal Medical Assistance Percentage Rates for Federal Matching Shares for Medicaid and Title IV–E Foster Care, Adoption Assistance and Guardianship Assistance Programs, 75 Fed. Reg. 66763.

Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association 82*(397), 171–185.

Ermini Leaf, D. (2017). Unified Method for Markov Chain Transition Model Estimation Using Incomplete Survey Data.

Ermini Leaf, D., B. Tysinger, D. P. Goldman, and D. N. Lakdawalla (2020). Predicting Quantity and Quality of Life with the Future Elderly Model. *Health Economics*.

Federal Bureau of Investigation (2020). Resource Guide Uniform Crime Reporting Program. `https://www.icpsr.umich.edu/web/pages/NACJD/guides/ucr.html`. [Online; accessed 18-July-2020].

Finkelstein, A. and N. Hendren (2020). Welfare Analysis Meets Causal Inference. *Journal of Economic Perspectives 34*(4), 146–67.

Finkelstein, A., E. F. Luttmer, and M. J. Notowidigdo (2013). What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption. *Journal of the European Economic Association 11*(Supplement 1), 221–258.

Finkelstein, A., E. F. P. Luttmer, and M. J. Notowidigdo (2009). Approaches to Estimating the Health State Dependence of the Utility Function. *American Economic Review 99*(2), 116–121.

García, J. L. and J. J. Heckman (2022). Three Criteria for Evaluating Social Programs. *Journal of Benefit-Cost Analysis 13*(3), 281–286.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2020). Quantifying the Life-Cycle Benefits of a Prototypical Early Childhood Program. *Journal of Political Economy 128*(7), 2502–2541.

Goldman, D. P., D. Ermini Leaf, J. Sullivan, and B. Tysinger (2018). The Future Elderly Model: Technical Documentation. Technical report.

Goldman, D. P., D. Ermini Leaf, and B. Tysinger (2018). The Future Adult Model: Technical Documentation. Technical report.

Grant, W. V. and C. G. Lind (1978). Digest of Education Statistics 1977-78.

Grant, W. V. and T. D. Snyder (1986). Digest of Education Statistics, 1985-86.

Greenland, S. and M. A. Mansournia (2015). Penalization, Bias Reduction, and Default Priors in Logistic and Related Categorical and Survival Regressions. *Statistics in Medicine 34*(23), 3133–3143.

Hansen, B. E. (2021). *Econometrics*. [Online; accessed 31-March-2021].

Heckman, J. J. and G. Karapakula (2021). Using a Satisficing Model of Experimenter Decision-Making to Guide Finite-Sample Inference for Compromised Experiments. *Econometrics Journal 24*(2), C1–C39.

Hendren, N. and B. Sprung-Keyser (2020). A Unified Welfare Analysis of Government Policies. *Quarterly Journal of Economics 135*(3), 1209–1318.

Hunt, P., J. Anderson, and J. Saunders (2017). The Price of Justice: New National and State-Level Estimates of the Judicial and Legal Costs of Crime to Taxpayers. *American Journal of Criminal Justice 42*(2), 231–254.

Hunt, P. E., J. Saunders, and B. Kilmer (2019). Estimates of Law Enforcement Costs by Crime Type for Benefit-Cost Analyses. *Journal of Benefit-Cost Analysis 10*(1), 95–123.

Imbens, G. W. and M. Kolesar (2016). Robust Standard Errors in Small Samples: Some Practical Advice. *Review of Economics and Statistics 98*(4), 701–712.

Institute for Social Research at the University of Michigan (2019). Panel Study of Income Dynamics, Public Use Dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan.

Kakalik, J. S., W. S. Furry, M. A. Thomas, and M. F. Carney (1981). *The Cost of Special Education: Summary of Study Findings*. Rand Corporation.

KFF (2012). Medicaid Financing: An Overview of the Federal Medicaid Matching Rate (FMAP). Technical report.

KFF's State Health Facts (2021). Federal Medical Assistance Percentage (FMAP) for Medicaid and Multiplier. Retrieved May 17, 2021. Technical report.

Liang, K.-Y. and S. L. Zeger (1986). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika 73*(1), 13–22.

McCollister, K. E., M. T. French, and H. Fang (2010). The Cost of Crime to Society: New Crime-Specific Estimates for Policy and Program Evaluation. *Drug and Alcohol Dependence 108*(1–2), 98–109.

Miller, T. R., M. A. Cohen, D. I. Swedler, B. Ali, and D. V. Hendrie (2020). Incidence and Costs of Personal and Property Crimes in the United States, 2017. *Available at SSRN 3514296*.

National Center for Education Statistics and United States. Office of Educational Research and Improvement. Center for Education Statistics and Institute of Education Sciences (US) (2000). *Digest of Education Statistics*. US Department of Health, Education, and Welfare, Education Division.

Quandt, R. E. (1958). The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes. *Journal of the American Statistical Association 53*(284), 873–880.

Quandt, R. E. (1972). A New Approach to Estimating Switching Regressions. *Journal of the American Statistical Association 67*(338), 306–310.

Rand, M. (2006). The National Crime Victimization Survey: 34 Years of Measuring Crime in the United States. *Statistical Journal of the United Nations Economic Commission for Europe 23*(4), 289–301.

RAND HRS Version P (2016). Produced by the RAND Center for the Study of Aging, With Funding From the National Institute on Aging and the Social Security Administration.

Shook-Sa, B., G. L. Couzens, and M. Berzofsky (2015). User's guide to national crime victimization survey (ncvs) direct variance estimation. *RTI International, Research Triangle Park, NC*.

Tax Policy Center (2020). Historical Average Federal Tax Rates for All Households. `https://www.taxpolicycenter.org/statistics/historical-average-federal-tax-rates-all-households`. [Online; accessed 14-January-2021].

United States Department of Justice (2010). National Judicial Reporting Program, 2006.

University of Michigan and the National Institute on Aging (2013). Health and Retirement Study, (2006 Biomarker Data) Sensitive Dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).

University of Michigan and the National Institute on Aging (2015a). Health and Retirement Study, (2010 Biomarker Data) Sensitive Dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).

University of Michigan and the National Institute on Aging (2015b). Health and Retirement Study, (2012 Biomarker Data) Sensitive Dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).

University of Michigan and the National Institute on Aging (2016). Health and Retirement Study, (RAND HRS Version P) Public Use Dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).

University of Michigan and the National Institute on Aging (2017). Health and Retirement Study, (2014 Biomarker Data) Sensitive Dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).

US Department of Justice (1984). The 1983 Jail Census.

US Department of Justice (1988). Report to the Nation on Crime and Justice.

US Department of Justice (1990). Census of Local Jails 1988.

US Department of Justice (1992). Census of State and Federal Correctional Facilities, 1990.