Health, Crime, and the Labor Market: Theory and Policy Analysis*

Yuki Otsu[†] C. Y. Kelvin Yuen[‡]

February 12, 2021

Abstract

Health has a significant impact on labor market outcomes, and thus on criminal decisions. We document that better health is associated with a lower probability of committing a crime. To study the economic mechanism behind this finding, we build an equilibrium search model of health, crime, and the labor market. We perform policy experiments in the model and study their impacts on crime and the labor market. The calibrated model shows that by introducing Medicare-for-all, the economy's crime rate would decrease by one percentage point while the aggregate output would increase by more than 10%.

JEL classification: E24, I10, I13, J01, J08

Keywords: health, crime, labor search, health insurance, crime policy

*The authors would like to thank Ian Fillmore, Barton Hamilton, Rody Manuelli, Yongseok Shin, and Ping Wang for their support and advice. The authors are also grateful to Naoki Aizawa, Scott Barkowski, Jacob Vogler, and participants at Asian Meeting of the Econometric Society in Seoul, China Meeting of the Econometric Society in Shanghai, Econometric Society European Meeting in Manchester, Midwest Macroeconomics Meetings in Athens, and internal seminars at WashU for helpful comments and discussions. We gratefully acknowledge support from Center for Research in Economics and Strategy (CRES) at Washington University in St. Louis.

[†]Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130. Email: y.otsu@wustl.edu

[‡]Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130. Email: kelvinyuen@wustl.edu

1 Introduction

How would the health status of an individual affect her criminal behavior? How would universal health insurance and other public policies affect crime and labor market outcomes? Labor market and criminal decisions are likely influenced by one's mental and physical health conditions, among other factors. It has long been documented that health has a significant effect on one's labor market outcomes.[1] For example, Cai (2009) and Jäckle and Himmler (2010) find that better health entails higher wages. Other researchers in the literature use various identification strategies and find similar results.[2] Health also has a considerable impact on other labor market outcomes. For example, poor health leads to a higher probability of being unemployed (García-Gómez, 2011; Halla and Zweimüller, 2013), longer non-employment duration (García-Gómez et al., 2010), and increased probability of non-participation (Trevisan and Zantomio, 2016). Healthier workers tend to have higher productivity resulting in higher wages. In particular, poor health hurts a worker's productivity because it is difficult to work efficiently when ill. Moreover, it may take unhealthy people longer to find a job than healthy people due to worse interview performance. Some of them may give up finding a job and transit out of the labor force.

Criminal activities can be driven by the worker's labor market performance. The seminal paper by Becker (1968) argues that criminal decisions are likely affected by the opportunity cost of committing a crime, which includes one's income and job prospects. Empirically, Edmark (2005) and Lin (2008) show that unemployment increases the risk of committing crimes. Similarly, Schnepel (2018) finds that employment is also an important factor in lowering recidivism. Yang (2017) shows that the labor market condition at the time of release significantly affects prisoners' reoffending behavior.

Given the paramount importance of health on the labor market outcomes, it is no wonder that health may significantly affect criminal activities. Individuals with relatively bad health conditions may find that lawful jobs may not be worth finding and choose to engage in illegal activities. Much

[1] The relationship between health and labor market performance is a large field in the literature and has a long history in labor economics. Currie and Madrian (1999) is an excellent survey of the research in early years. See Appendix A for more evidence using the PSID data.

[2] For example, García-Gómez et al. (2013) find that, using acute hospital admissions as health shocks, bad health shocks lower earnings. If health status affects earnings, medical treatment could improve it. Stephens and Toohey (2018) investigate the impact of treatment to reduce the risk of coronary heart disease on labor market outcomes and find that the treatment increases earnings and family income.

of the empirical literature, however, focuses on the relationship between mental illness and criminal behaviors.[3] Indeed, while there is evidence that criminals are less healthy than non-criminals,[4] direct evidence of the effect of physical health on crime is limited. One exception is Schroeder et al. (2011), who find that a decline in physical health is associated with increasing the probability of committing crimes. There is literature exploring health care as a tool to improve offenders' re-entry into society.[5] Wen et al. (2014) and Wen et al. (2017) show substance use disorder (SUD) treatment reduces both violent and property crime rates. Cuellar et al. (2004) and Cuellar et al. (2006) find that SUD treatment and mental health treatment detains youth crime. Although these papers study the effect of health (or medical treatment) on crime, the mechanism is not clear.

In this paper, we study empirically and theoretically how general health status affects an individual's criminal activities. While the effect of health on labor market outcomes has been well-studied in the literature, studies of the effect of health on criminal behaviors, especially through the labor market channel, have been rare. A unified model of health, crime, and the labor market is essential to answer the question. The unified model allows us to perform a wide range of policy experiments to evaluate their impacts on both the labor market and criminal activities.

Empirically, we use the microdata from the National Longitudinal Survey of Youth 1997 (NLSY97) to investigate the relationship between health and criminal activity. We find that on average better health is associated with a lower probability of committing a crime. This is true for both property crime and violent crime and is robust to a wide range of specifications. For example, the average elasticity of property crime of an individual improving the health status from Poor to Good is -0.82. This translates to an average marginal effect of about -2.30 percentage points per year as the individual's health improves from Poor to Good.

To understand the mechanism behind the empirical relationship and to study its policy implications, we build an equilibrium model of health, crime, and the labor market. We incorporate health and crime into the standard search and matching model (Mortensen and Pissarides, 1994). There are four labor market states in the model: employed, unemployed, not-in-labor-force, and in prison.

[3] In Sweden, one in twenty violent crimes is committed by patients with serious mental illness (Fazel and Grann, 2006). Moreover, it is well known that health status and recidivism have a negative correlation. In particular, there is evidence that mental health status is associated with violent behavior (Fazel et al., 2009).

[4] For instance, Fazel and Baillargeon (2011) and Wilper et al. (2009) find that US prisoners are in general less healthy than the general population in the US.

[5] For more recent research, see, e.g., Regenstein and Rosenbaum (2014), Kinner and Wang (2014), Aalsma et al. (2015), Frank and McGuire (2010), Marcotte and Markowitz (2010), and references therein.

The criminal decisions are modeled á la Engelhardt et al. (2008), where individuals may have an opportunity to commit a crime in each period. Also, individuals in the model are heterogeneous in health, ability, and health insurance coverage. Health insurance status is modeled stochastically as in Aizawa and Fang (2020). In the model, health affects the productivity of the workers positively and the job search cost negatively. To capture the labor force participation margin in the data, we explicitly allow endogenous participation decisions. Health insurance affects both the transitions of health status and the production elasticity of health and ability. We find in the model that healthier individuals are more likely to participate in the labor market and that they are more likely to be employed, consistent with the data. We also find that individuals with better health status or higher ability would be less likely to commit a crime, since they generally have better labor market outcomes, and thus have a higher opportunity cost of committing a crime.

The model is then calibrated to the United States economy. We show quantitatively that both the health status and health insurance have a considerable impact on the labor market. In general, healthier workers face a higher labor force participation rate and a higher employment rate. Also, we find that criminal activity is negatively related to the worker's health status and education level, and is more so for those with health insurance. For uninsured workers, the crime rate for those with the best health and education is two percentage points lower than that for those with the worst health and education. For those with health insurance, the difference is more striking. The crime rate of insured workers with the best health status and education is only one-sixth of those with the worst health and education. The crime rate for the insured individuals is more elastic to health and education mainly because uninsured individuals are much less likely to participate in the labor market, and are less sensitive to the labor market effects of health and education.

The model developed allows us to perform a wide range of policy experiments. We first evaluate the impact of health insurance coverage on the labor market outcomes and criminal decisions. We find that health insurance policies such as Medicare-for-all and the Employer Mandate under the Affordable Care Act (ACA) would increase the participation rate and the aggregate output. However, they have different effects on the crime rates of the economy. For example, our model estimates that Medicare-for-all would lead to more than one percentage point decrease in the crime rate. On the other hand, if all jobs are insured by employer-sponsored health insurance, crime rates would increase by about one percentage point. We also find that, compared to the Employer

Mandate policy, Medicare-for-all would reduce income inequality more significantly (−21.6% vs. +0.1%) and have more improvement on welfare (+8.3% vs. +6.7%). Other public policies are also investigated. Consistent with the literature, an increase in unemployment benefits would lead to a decrease in the participation rate and employment-population ratio across individuals with all health statuses. With harsher sentencing policy, the crime and unemployment rates increase, while the participation rate decreases. However, the quantitative effects of these impacts depend on one's health and health insurance status.

This paper contributes to the literature in three ways. First, although a large body of literature emphasizes mental health, general health does not gather much attention so far. We show empirically that general health status significantly affects criminal behaviors using a nationally representative sample.[6] Second, the model developed in this paper illustrates the importance of the labor market channel in explaining the effects of health on crime, which has not been discussed in the literature. Third, we perform numerous policy analyses in a model with heterogeneity in health and health insurance status. While the effects of the public policies are generally found to be monotonic in the literature, we find that they depend on an individual's health status and health insurance coverage in many cases. These differential impacts show that such a comprehensive model is essential when conducting a public policy analysis. To the best of our knowledge, this paper is the first attempt to construct such a framework.

Relevant literature Many papers in the literature use a search model to study the relationship between crime and the labor market. For example, Burdett et al. (2003) and Burdett et al. (2004) study wage dispersion in the labor market in a search model with criminal behavior. The model is also employed by Engelhardt et al. (2008) and Otsu (2016), who study the effects of a crime option in the labor market. However, all the papers mentioned above assume homogeneity of workers and do not allow participation decisions in the model. Participation decision in a search and matching framework is introduced, for example, by Garibaldi and Wasmer (2005). In their model, individuals choose whether to participate in the labor market when they are not employed. However, they do not allow criminal decisions and heterogeneity in health.

[6] In NLSY97, the correlation between having good general health status (Good, Very good, Excellent) and low depression (None, Some) is only 11% in 2000. This shows that the general health status is not only a proxy for the mental health status.

This paper also relates to recent literature examining the impact of health policy on the labor market. Aizawa and Fang (2020) and Nakajima and Tuzemen (2017) study the effects of the ACA in an equilibrium job search model. Also, Dey and Flinn (2005) analyze an employer-provided health insurance contract under a search framework and explore its effects on job mobility and welfare. They assume a pair of wage and health insurance as a contract and find that job mobility is not affected by the health insurance provision. Finally, Pashchenko and Porapakkarm (2013) evaluate the impact of health insurance reform in an equilibrium life-cycle model. By including criminal decisions and an explicit production function, our model can evaluate the impact of health insurance policy on the crime rate and aggregate output.

The remainder of the paper is organized as follows. In Section 2, we empirically study the relationship between an individual's health status and criminal activities. We develop empirical evidence in the section. In Section 3, we build a model that incorporates health and crime into a search and matching framework. We quantitatively study the impact of health on both the labor market and criminal decisions in Section 4. In Section 5, we perform policy analysis in the model. Section 6 concludes the paper.

2 The empirical relationship between health and crime

This section investigates the empirical relationship between an individual's health status and criminal behaviors in the United States using microdata from the National Longitudinal Survey of Youth 1997 (NLSY97). To this end, we estimate a regression model of individual's criminal behavior on their health status and demographic characteristics. In Appendix A, we also document the relationship between health and labor market outcomes using the Panel Study of Income Dynamics (PSID).[7]

2.1 Data

NLSY97 is a nationally representative longitudinal sample of about 9,000 individuals who were 12 to 16 years old as of 1996. We use the annual data from 1998 to 2011, after which the survey started to be conducted biannually. This period is appropriate to analyze criminal behavior because it

[7] We find that better health is associated with a higher probability of employment and higher wage, controlling for other important factors.

covers youth and young adults who are in the crime-prone period of their life. The survey contains self-reported information about individual's criminal behaviors.[8]

We define two dummy variables for violent behavior and property crime as follows. A dummy variable *Violent behavior* equals one if an individual has attacked someone since the last interview and equals zero otherwise.[9] Also, a dummy variable *Property crime* equals one if an individual has stolen any money or committed other property crime.[10] For the measure of individual's health, a vector of dummy variables $\mathbf{HealthDummy} = [Fair, Good, Very\ Good, Excellent]$ is derived from a question asking the general health status of an individual.[11][12]

Table A.1 summarizes the detailed data selection process, and the summary statistics of our sample are in Table A.2. On average, we see that healthier individuals are less likely to engage in violent behavior or commit property crimes.

2.2 Crime regression

We perform a regression analysis to study the relationship between health status and criminal behavior. Our specification of the regression model is as follows:[13]

$$\Pr\left(crime_{it+1} = 1 | \mathbf{HealthDummy}_{it}, \mathbf{X}_{it}^{c}, \eta_i, \eta_t\right) = G\left(\beta_0 + \boldsymbol{\beta}_1 \mathbf{HealthDummy}_{it} + \boldsymbol{\beta}_X \mathbf{X}_{it}^{c} + \eta_i + \eta_t\right) \quad (1)$$

[8] The two measures of crime are self-reported. As is common with other survey data about crime, criminal behaviors are likely underreported.

[9] The relevant question in the survey is: "Have you ever attacked someone with the idea of seriously hurting them or have a situation end up in a serious fight or assault of some kind?"

[10] The relevant questions are: "How many times have you stolen something from a store, person or house, or something that did not belong to you worth 50 dollars or more including stealing a car in the last 12 months?", "How many times have you stolen something from a store or something that did not belong to you worth less than 50 dollars since the last interview on [date of last interview]?" and "How many times have you committed other property crimes in the last 12 months?"

The variable *Property crime* equals one if the sum of the answers to all of the above question is non-zero and equals zero otherwise. In the data, among those who have committed at least one of the property crimes, about 30% have stolen something more than $50 at least once.

[11] NLSY97 measures general health status from Excellent to Poor. The relevant question in the survey is, "Now, I'd like to ask you some questions about your general state of health. In general, how is your health?"

[12] The self-reported health status is inherently subjective. Recently, Hosseini et al. (2018) construct a frailty index using more objective health symptoms. While there are important differences when studying health dynamics over the life cycle, the two measures tend to be highly correlated in the cross-sectional data.

[13] We use the crime variable from the year $t+1$ rather than t so that the timing of the crime variable and other variables are consistent. In NLSY97, health status is asked at the time of the interview, while the crime variable is derived from crimes committed since the last interview.

where $crime_{it+1}$ is a crime dummy of an individual i in survey year $t+1$, which is either a dummy variable for property crime or violent behavior. The function G is the logistic function. $\mathbf{HealthDummy}_{it}$ is a vector of health status dummies, and we set Poor as the base level. $\mathbf{X}^c_{it}$ contains control variables, which include a female dummy, race dummies (black, Hispanic, and mixed), age, squared age, years of education, and marital status. Finally, η_i and η_t are respectively the individual and year fixed effects.

Table 1 shows the regression coefficients for property crimes. The first three columns are based on a logit model, and the coefficients represent the marginal effects on the log-odds ratio. The first column contains the result from the pooled logit regression, and the second column contains that from the logit regression with individual fixed effects. Both columns show that health status has a mildly negative correlation with property crimes, and the magnitude becomes larger for better health. Table B.9 in Appendix B shows the average elasticity of the probability of crime with respect to health derived from the logit regression models. For example, the average semi-elasticity of property crime of an individual improving the health status from Poor to Good is -0.82 after controlling for individual fixed effects. Given the average property crime rate is about 2.81%, this would entail an average marginal effect of about -2.30 percentage points. Figure 1 shows the predicted probability of committing a property crime with different health status using the logit model specification in column (1). The predicted probability decreases monotonically as health improves. Since column (2) includes individual fixed effects, the coefficients are derived from the conditional logit estimation. Therefore, a smaller sample size is used in column (2), as the conditional logit estimation requires outcome variation within an individual. We can see that the result is similar after we add the individual fixed effects to the model. In column (3), we aggregate a dummy for Good, Very Good, and Excellent into one variable and use it as a different measure of health. The result also shows the negative correlation between health and crime.

Note that there could potentially be an endogeneity issue in the crime regressions. For example, poverty could be a source of both worse health and more crime. To mitigate the endogeneity issue, we use difference-in-differences (DID) and system generalized method of moments (SGMM) specifications. In the DID column in the regression results, we measure the effect of a negative health shock between $t-1$ and t on outcomes at $t+1$. The approach is an extension of a difference-in-differences approach used in García-Gómez (2011). A negative health shock is the change of health status from good health (Good, Very Good, Excellent) to bad health (Poor, Fair). The estimated equation is

$$y_{it+1} = \alpha_0 + \alpha_1 NegativeHealthShock_{it} + \alpha_2 GoodHealth_{it-1} + \alpha_X X_{it} + \eta_i + \eta_t + \epsilon_{it}. \quad (2)$$

$NegativeHealthShock_{it}$ takes one if the health history is (Good,Bad) and zero otherwise. Possible health histories at $t-1$ and t are (Good,Good), (Good,Bad), (Bad,Good), and (Bad,Bad). To focus on a negative health shock, We exclude the sample if it has bad health at both $t-1$ and t: (Bad,Bad). The equation has the term $Goodhealth_{it-1}$ to distinguish (Bad,Good) from other two cases. Hence, $NegativeHealthShock_{it}$ captures the difference between (Good,Good) and (Good,Bad). The DID result indicates a negative health shock defined by a negative change in health status increases the probability of committing a property crime by one percentage point compared with those who do not experience the shock.

Alternatively, we estimate the relationship between current health status and criminal behavior in a dynamic panel model. In the baseline specification, we have used the forward variables of the outcomes one period ahead. However, the regression model ignores the dynamic interaction between health and crime, and it is well-documented that the estimates of a dynamic panel model with fixed effects are inconsistent. The SGMM addresses this issue by controling for the past values of the outcome variable and taking advantage of the levels and difference equations. Let y be an outcome variable. We consider the following model:

$$y_{it} = \gamma_1 y_{it-1} + \gamma_2 y_{it-2} + \beta GoodHealth_{it} + \alpha_{en} X_{it}^{en} + \alpha_{ex} X_{it}^{ex} + \eta_i + \eta_t + \epsilon_{it}. \quad (3)$$

The current value of an outcome depends on its past values (due to persistence) and health status at t as well as other covariates X_{it}^{en} and X_{it}^{ex}. This model allows endogeneity between health and outcomes.[14] The result from the SGMM model also supports that better health leads to a lower propensity to commit a crime. These results all suggest that healthier individuals are less likely to commit property crimes.

Table 2 shows the regression results for violent behavior. Similar to property crimes, the result

[14] See Appendix B for more details.

Table 1: Property crime regression on health status (NLSY97)

Property crime	(1) Logit	(2) Logit	(3) Logit	(4) DID	(5) SGMM
Fair	-0.441	-0.654			
	(0.227)	(0.302)			
Good	-0.802	-0.845			
	(0.219)	(0.299)			
Very Good	-0.869	-0.869			
	(0.218)	(0.303)			
Excellent	-1.104	-0.738			
	(0.219)	(0.306)			
Good Health			-0.229		-0.0433
			(0.108)		(0.0252)
Negative Health Shock				0.00997	
				(0.00415)	
Control	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Individual FE	N	Y	Y	Y	Y
Observations	82,776	14,391	14,391	69,409	61,957
Number of individuals		1,409	1,409	8,253	7,980

Note: Robust standard errors in parentheses. The dependent variable *property crime* is made from questions that ask if stealing money more/less than 50 dollars or if committing other property crime. Year FE for logit models are two-year dummies. Control variables are a female dummy, race dummies, age, age squared, years of education, and marital status. See Appendix B for details about SGMM.

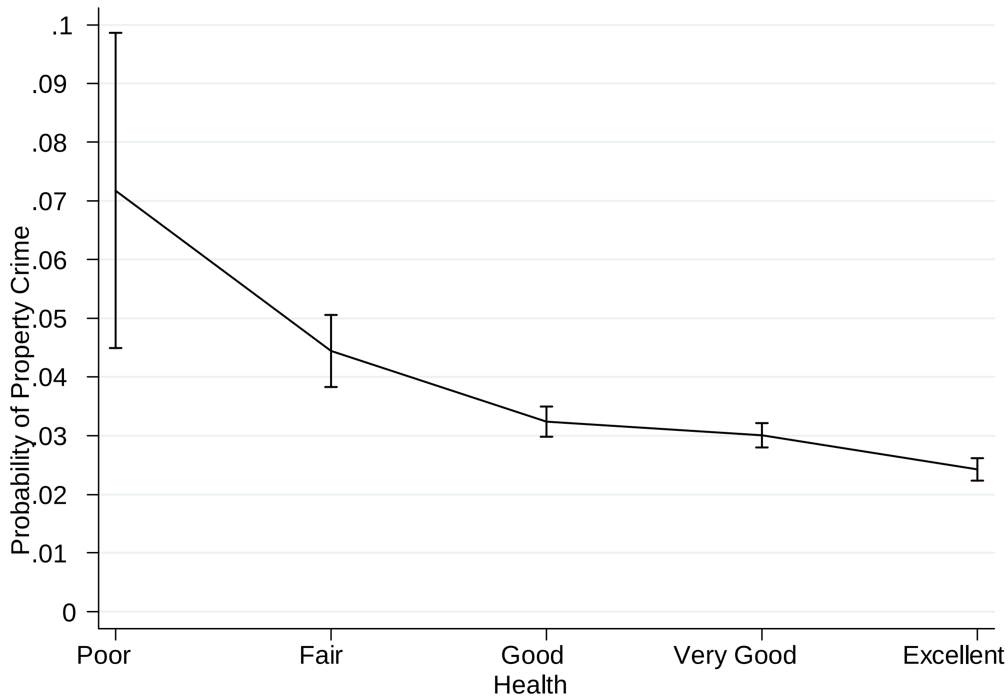

Figure 1: Predicted crime rate

Notes: The predicted probability is based on the logit model (1) using the NLSY97 data. The vertical lines at each health status show 95% confidence interval of the estimates. See Section 2 for the regression models and data construction.

Table 2: **Violent crime regression on health status (NLSY97)**

Violent behavior	(1) Logit	(2) Logit	(3) Logit	(4) DID	(5) SGMM
Fair	-0.737	-0.544			
	(0.189)	(0.245)			
Good	-1.056	-0.593			
	(0.181)	(0.243)			
Very Good	-1.427	-0.742			
	(0.182)	(0.247)			
Excellent	-1.424	-0.591			
	(0.182)	(0.250)			
Good Health			-0.136		-0.0486
			(0.0940)		(0.0285)
Negative Health Shock				0.00646	
				(0.00423)	
Control	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Individual FE	N	Y	Y	Y	Y
Observations	82,776	14,047	14,047	69,409	61,968
Number of individuals		1,404	1,404	8,253	7,979

Note: Robust standard errors in parentheses. Violent behavior is defined from the question "Have you attacked someone with the idea of seriously hurting them or have had a situation end up in a serious fight or assault of some kind?" Year FE for logit models are two-year dummies. Control variables are a female dummy, race dummies, age, age squared, years of education, and marital status. See Appendix B for details about SGMM.

shows that health has a negative relationship with violent behaviors. In this case, the average semi-elasticity of violent crime of an individual improving the health status from Poor to Good is -0.58 after controling for individual fixed effects.

The same endogeneity concern applies to violent behavior, so we use DID and SGMM from columns (4) and (5). The DID result indicates that a negative health shock increases the probability of committing a violent crime by 0.6 percentage point. Although the coefficient is not statistically significant, the magnitude is smaller than the impact for property crime. The point estimator in the SGMM model also shows a negative relationship between health and violent behavior.

In summary, the empirical results support that better health is associated with fewer criminal activities, especially for property crimes. In Appendix A, we check the robustness of these results by adding additional controls: employment status, drug usage, health insurance coverage, and income levels. However, none of these specifications would change the significance of the association between health and crime. One potential channel is that better health leads to better employment outcomes, which reduces criminal behaviors. In the following section, we build an equilibrium model that incorporates health, crime, and the labor market to explain the mechanism behind these findings.

3 A model of health, crime, and the labor market

In this section, we present an equilibrium model of health, crime, and the labor market. We extend the standard search and matching model of the labor market (Mortensen and Pissarides, 1994) with criminal behavior and health heterogeneity.[15] We model the criminal decisions á la Engelhardt et al. (2008). We also include elements such as health status and an out-of-labor-force state into the model, which would be essential to our analysis. Health insurance status is modeled stochastically, as in Aizawa and Fang (2020). In Appendix D, we lay out more details to close the model.

3.1 Workers

Workers are infinitely lived and risk-neutral with common discount factor β. In each period of their life, they belong to one of the four states in the labor market: employed (W), unemployed (U), not-in-labor-force (N), and in prison (P). A worker can freely move between unemployed, if she is searching for a job, and not-in-labor-force if she is not. U and N are grouped together as non-employment (NE). Figure 2 shows the flow diagram of the states.

Workers are heterogeneous in health status h, ability y, and health insurance status ω. We consider y to be time-invariant and include such attributes as innate ability and schooling, which would enhance a worker's productivity. For simplicity, we assume h is continuous with support $[h_{min}, h_{max}]$, and y is discrete with support $\{y_1, ..., y_J\}$. Unemployed workers and vacant firms are searching randomly in segmented markets for each y. The reason we assume segmented markets for each y is that jobs generally have different education or experience requirements. Upon matching

[15] We consider health status in the model as the general health that affects the productivity of working but not that of crime.

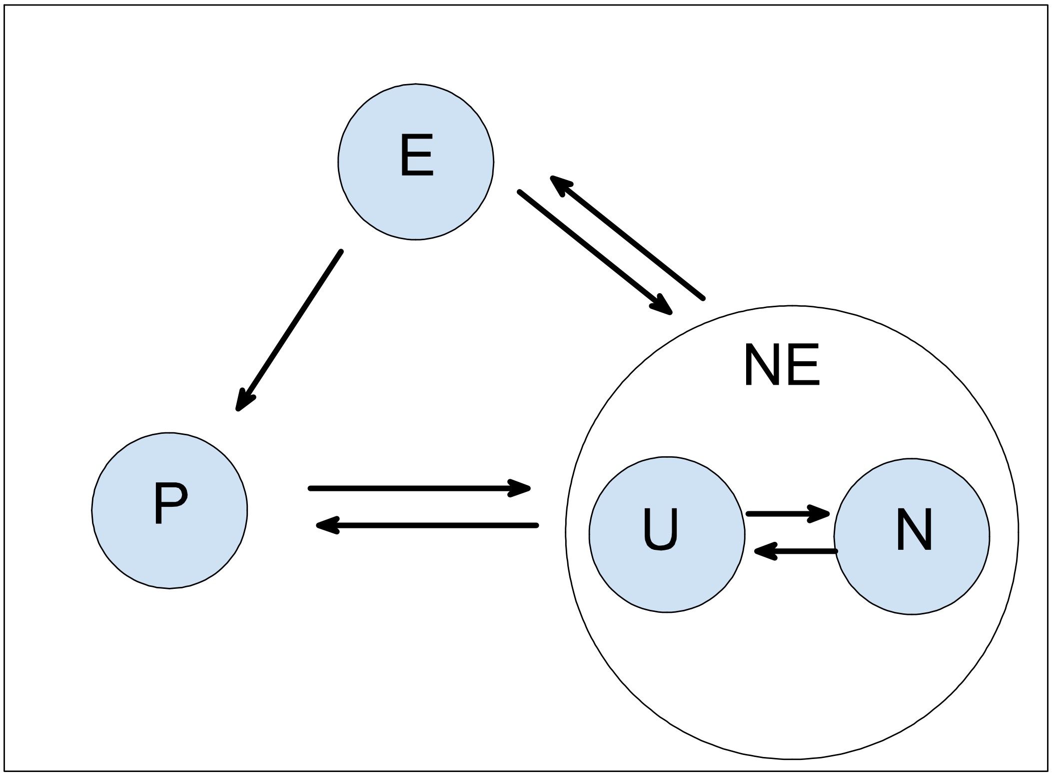

Figure 2: Labor market flows

with a firm, the employment pair jointly produces a flow value of output $\phi(h, y; \omega)$, which is strictly increasing in h and y. Finally, workers can be uninsured, insured by public health insurance (PHI), or insured by employer-sponsored health insurance (ESHI). Hence, we denote $\omega \in \{0, 1, 2\}$ as the health insurance status as follows.

$$\omega = \begin{cases} 0 & \text{if the individual is uninsured} \\ 1 & \text{if the individual is insured by public health insurance} \\ 2 & \text{if the individual is insured by employer-sponsored health insurance} \end{cases}$$

3.1.1 Criminal opportunity

Any person, except those who are in prison, has some chance to commit a crime in each period.[16] For each state $X \in \{U, W, N\}$, the probability of receiving a criminal opportunity is λ. The criminal

[16] The assumption that prisoners have no chance to commit a crime is mainly technical. In this model, prisoners are released stochastically with a constant release rate. As a result, there is essentially no punishment if they are allowed to commit a crime.

opportunity m is drawn from the distribution function $H(m)$ with support $[m_{\min}, m_{\max}]$, [17] where $m_{\min} \geq 0$. After the realization of m, the worker can choose whether or not to commit the crime. If not, then she would receive $\tilde{X}$, the continuation value of the state X which will be discussed in the following subsections. On the other hand, if she chooses to commit the crime, she will earn m, and there is a probability π that she will get caught and be sent to prison in the next period.[18] Finally, let $\tilde{V}^P(h, y)$ be the value function of a worker in prison with health h and ability y.

Therefore, the value of potential criminal opportunity for a worker in state X, with health h and ability y is given by

$$K^X(h, y; \omega) = \int_{m_{\min}}^{m_{\max}} \max\left\{m + \beta\left[\pi\tilde{V}^P(h, y) + (1-\pi)\tilde{V}^X(h, y; \omega)\right], \beta\tilde{V}^X(h, y; \omega)\right\} dH(m) \tag{4}$$

3.1.2 Unemployed workers

Unemployed workers receive a flow utility b of non-employment activity, which may include such things as leisure, home production, and unemployment benefits. They face a search cost $\sigma(h)$ when searching for a job. We assume $\sigma'(h) < 0$ and $\sigma''(h) > 0$. Therefore, workers with better health face lower search costs. This captures the fact that healthier workers are more active in a job search. Lastly, the expected loss from crime is τ^c, which is taken as given by individuals and is determined in equilibrium.[19] Therefore, the value of an unemployed worker with health h and ability y is given by

$$V^U(h, y; \omega) = b - \sigma(h) - \tau^c + \lambda K^U(h, y; \omega) + (1-\lambda)\beta\tilde{V}^U(h, y; \omega) \tag{5}$$

where $\tilde{U}(h, y; \omega)$ is the continuation value given by

$$\tilde{V}^U(h, y; \omega) = \int_{h_{\min}}^{h_{\max}} \varphi^U_{PHI}\tilde{U}(h', y; 1) + \left(1-\varphi^U_{PHI}\right)\tilde{U}(h', y; 0)\, dF^\omega(h'|h) \tag{6}$$

$$\tilde{U}(h', y; \omega') = \rho_y \begin{bmatrix} \varphi_{ESHI}\max\left\{V^W(h', y; 2), V^{NE}(h', y; \omega')\right\} \\ +(1-\varphi_{ESHI})\max\left\{V^W(h', y; \omega'), V^{NE}(h', y; \omega')\right\} \end{bmatrix} + \left(1-\rho_y\right)V^{NE}(h', y; \omega') \tag{7}$$

[17] We assume health has no direct impact on the realization of the value of crime, since it is not clear empirically how it would be affected by health. In this paper, we focus on the indirect impact of health status on criminal behavior through the labor market performance.

[18] As in Engelhardt et al. (2008), we assume all individuals have the same apprehension rate π. While it is natural to think that π may depend on health and ability, there is insufficient data to identify the relationship.

[19] See Appendix D for the details. In the model, crimes are committed randomly to the individuals in the economy. Hence, the loss from crime is constant across individuals.

Thus, in the next period, the new health status h' is drawn from the conditional distribution $F^{\omega}(h'|h)$ which satisfies the first-order stochastic dominance property[20]

$$F^{\omega}\left(h'|h_1\right) \leq F^{\omega}\left(h'|h_2\right) \tag{8}$$

for all $h_1 \geq h_2$ in the support of h. There is a probability φ^U_{PHI} that the worker will be insured by public health insurance in the next period. Then for each realized health insurance status, an unemployed worker will meet with a firm with the rate ρ_y, which is determined endogenously.

Moreover, unemployed workers meet a job with employer-sponsored health insurance with probability φ_{ESHI}. Otherwise, the worker stays with her health insurance status drawn previously. Note that in this setting, if the individual draws both PHI and ESHI, she would take ESHI. In any case, if the value of working V^W is higher than that of non-employment V^{NE}, which can be either unemployment or not-in-labor-force:

$$V^{NE}(h, y; \omega) = \max\left\{V^U(h, y; \omega), V^N(h, y; \omega)\right\} \tag{9}$$

for $\omega \in \{0, 1\}$, then the worker will be employed.

3.1.3 Employed workers

If matched, employed workers earn a wage $w(h, y; \omega)$ which will be determined by Nash bargaining. Hence, the value of being employed is given by

$$V^W(h, y; \omega) = w(h, y; \omega) - \tau^c + \lambda K^W(h, y; \omega) + (1 - \lambda)\beta \tilde{V}^W(h, y; \omega) \tag{10}$$

ESHI will remain for the rest of the career. Those who do not have one would have to draw public health insurance with probability φ^W_{PHI}. There is an exogenous separation rate s. In addition, the worker may be endogenously separated from the firm if the value of being employed is smaller than

[20] This property says that a healthier individual in the current period will on average be healthier in the next period as well.

that of non-employment. Therefore, the continuation value of the employed worker is

$$\tilde{V}^W(h,y;\omega) = \begin{cases} \int_{h_{\min}}^{h_{\max}} \tilde{W}(h',y;2)\, dF^{\omega}(h'|h) & \text{when } \omega = 2 \\ \int_{h_{\min}}^{h_{\max}} \varphi_{PHI}^{W} \tilde{W}(h',y;1) + \left(1-\varphi_{PHI}^{W}\right) \tilde{W}(h',y;0)\, dF^{\omega}(h'|h) & \text{when } \omega \in \{0,1\} \end{cases} \tag{11}$$

$$\tilde{W}(h',y;\omega') = \left[(1-s)\max\left\{V^W(h',y;\omega'), V^{NE}(h',y;\omega')\right\} + sV^{NE}(h',y;\omega')\right] \tag{12}$$

For $\omega = 2$, we define $V^{NE}(h,y;2) = V^{NE}(h,y;1)$ for the sake of notational simplicity.

3.1.4 Non-participants

A worker who is not in the labor force still receives the flow utility b. However, since the worker is not searching for a job, there is no search cost involved. Therefore, the value of a worker not in the labor force is given by

$$V^N(h,y;\omega) = b - \tau^c + \lambda K^N(h,y;\omega) + (1-\lambda)\beta\tilde{V}^N(h,y;\omega) \tag{13}$$

In the next period, the worker may draw public health insurance with probability φ_{PHI}^{N}, and choose to stay out of the labor force or participate in the labor market and look for a job. Hence, the continuation value is given by

$$\tilde{V}^N(h,y;\omega) = \int_{h_{\min}}^{h_{\max}} \varphi_{PHI}^{N} V^{NE}(h',y;1) + \left(1-\varphi_{PHI}^{N}\right) V^{NE}(h',y;0)\, dF^{\omega}(h'|h) \tag{14}$$

3.1.5 Prisoners

Prisoners receive a flow utility $x < b$. They are still subject to the expected criminal loss.[21] Also, prisoners are always insured by public health insurance. In the next period, there is a probability δ that a prisoner is released to the pool of non-employment. Otherwise, the prisoner stays in prison. There is a probability φ_{PHI}^{P} that they will be insured by public health insurance upon release from prison. Hence, the value of being in prison would be

$$V^P(h,y) = x - \tau^c + \beta\tilde{V}^P(h,y) \tag{15}$$

$$\tilde{V}^P(h,y) = \int_{h_{\min}}^{h_{\max}} \begin{bmatrix} \delta\left[\varphi_{PHI}^{P} V^{NE}(h',y,1) + \left(1-\varphi_{PHI}^{P}\right) V^{NE}(h',y,0)\right] \\ +(1-\delta)V^P(h',y) \end{bmatrix} dF^1(h'|h) \tag{16}$$

[21] For the sake of simplicity, we assume that criminal loss also applies to prisoners. It can be shown that both qualitative and quantitative results are similar if prisoners don't suffer from criminal loss.

3.2 Criminal and labor market decisions

Note from (4) that after the realization of the value of a crime m, the value of committing the crime is strictly increasing with the value of m, whereas the value of not doing so is constant with respect to m. Therefore, the criminal decision can be characterized by a cutoff value $\bar{m}$, above which the individual would choose to commit the crime. By rearranging (4), we have

$$K^X(h,y;\omega) = \beta\tilde{V}^X(h,y;\omega) + \int_{\bar{m}^X(h,y)}^{m_{\max}} \left(m - \bar{m}^X(h,y;\omega)\right) dH(m) \tag{17}$$

where

$$\bar{m}^X(h,y;\omega) = \beta\pi\left[\tilde{V}^X(h,y;\omega) - \tilde{P}(h,y;\omega)\right]. \tag{18}$$

For analytical purposes, we assume that the parameters satisfy some mild regularity conditions, shown to hold in the baseline calibration in Section 4.[22] Then we have the following proposition regarding the criminal decision of individuals:

Proposition 1 *Suppose* $\varphi^W_{PHI} = \varphi^U_{PHI} = \varphi^N_{PHI}$ *and* $1 - s \geq \rho_y$*, then for each* h, y, ω *in their support,*

(i) The cutoff crime value is the highest for the employed workers, followed by the unemployed and non-participants. i.e.

$$\bar{m}^W(h,y;\omega) \geq \bar{m}^U(h,y;\omega) \geq \bar{m}^N(h,y;\omega) \geq 0 \tag{19}$$

and

(ii) For $X \in \{U,W,N\}$*,* $\bar{m}^X(h,y)$ *is increasing in both the health status and the ability,*

$$\frac{\partial \bar{m}^X(h,y;\omega)}{\partial h} \geq 0 \tag{20}$$

$$\frac{\partial \bar{m}^X(h,y;\omega)}{\partial y} \geq 0 \tag{21}$$

Proof. *See Appendix C.* ■

The first part of the proposition states that, other things equal, employed workers are less likely to commit a crime than unemployed workers, which are less likely to commit a crime than those

[22] The full set of conditions is listed in Appendix C.

who are not in the labor force. The second part concerns health and ability: workers with better health status or a higher ability are less likely to commit a crime, given the same labor force status.

The results are intuitive and stem from the insights of Becker (1968) that criminal decisions are likely affected by the opportunity cost of committing a crime. Being employed in a job certainly has a higher opportunity cost of committing crimes than being unemployed since the criminal may lose the job if getting caught.[23] Unemployed workers are less prone to criminal activity than non-participants since it takes one period to participate in the labor market in the model. Moreover, workers with better health or better ability tend to have better labor market outcomes and tend to have a lower propensity to commit a crime.

The labor market decisions of the workers can also be characterized by cutoff strategies. In the case of interior cutoffs, the threshold values for employment decisions can be determined by

$$V^{U}\left(\bar{h}_{LFP}, y; \omega\right) = V^{N}\left(\bar{h}_{LFP}, y; \omega\right) \tag{22}$$

$$V^{W}\left(\bar{h}_{EMP}, y; \omega\right) = V^{NE}\left(\bar{h}_{EMP}, y; \omega'\right) \tag{23}$$

where $\bar{h}_{LFP}(y;\omega)$ denotes the cutoff health status to participate in the labor force, and $\bar{h}_{EMP}(y;\omega,\omega')$ denotes the cutoff health status to be employed.[24] Therefore, individuals with health better than $\bar{h}_{LFP}(y;\omega)$ would participate in the labor market, and those with health better than $\bar{h}_{LFP}(y;\omega)$ would be employable in the sense that once they match with firms, the total surplus would be positive.

We do not assume that the interior threshold values of (22) and (23) exist. Indeed, we might have a corner solution. For example, we set $\bar{h}_{LFP}(y;\omega) = h_{\min}$ if every individual with ability y chooses to participate in the labor market. On the other hand, $\bar{h}_{LFP}(y;\omega) = h_{\max}$ if no one with ability y would participate. However, we assume that the curvature of the value functions ensures that the threshold values would be unique.[25] Then we have the following

Proposition 2 *Suppose* $\varphi^{W}_{PHI} = \varphi^{U}_{PHI} = \varphi^{N}_{PHI}$ *and* $1 - s \geq \rho_y$*, then for each* y, ω *in their support,*

[23] The condition $1 - s \geq \rho_y$ ensures that the continuation value of being employed is weakly larger than that of being unemployed.

[24] Since the Nash bargaining solution is always privately efficient, this is also the cutoff health above which the match has a positive surplus.

[25] Technically, the curvature of the value functions with respect to the health status depends on that of the production function $\phi(h, y; \omega)$ and search cost function $\sigma(h)$. This condition is satisfied quantitatively under the functional forms specified in the quantitative exercise.

the cutoff values $\bar{h}_{LFP}$ and $\bar{h}_{EMP}$ are decreasing in the ability y, i.e.

$$\frac{\partial \bar{h}_{LFP}(y;\omega)}{\partial y} \leq 0 \tag{24}$$

$$\frac{\partial \bar{h}_{EMP}(y;\omega,\omega')}{\partial y} \leq 0 \tag{25}$$

Proof. *See Appendix C.* ■

Intuitively, since the product $\phi(h,y;\omega)$ is increasing in both h and y, all value functions are also increasing in both variables. As a result, a worker with higher ability y would be more likely to participate in the labor market (i.e., lower $\bar{h}_{LFP}$), and more likely to be employed if matched (i.e., lower $\bar{h}_{EMP}$).

3.3 Firms

3.3.1 Free entry

Firms post vacancies with a flow cost γ. With endogenous probability q_y, a vacant firm meets with an unemployed worker with ability y. Let $\zeta_y(h)$ be the probability density that the firm meets with a worker with health h in the y market. Upon meeting with the worker, the firm chooses whether to hire her depending on the value of matching and the health insurance status. Hence, free entry of firms implies that, for any y,

$$\gamma = \beta q_y \int_{h_{\min}}^{h_{\max}} \zeta_y(h) \left[\begin{array}{c} \varphi_{ESHI} \max\{\Pi(h,y,2),0\} \\ +(1-\varphi_{ESHI}) \left(\begin{array}{c} \varphi_{PHI}^{U} \max\{\Pi(h,y;1),0\} \\ +\left(1-\varphi_{PHI}^{U}\right) \max\{\Pi(h,y;0),0\} \end{array} \right) \end{array} \right] dh \tag{26}$$

where $\Pi(h,y;\omega)$ is the value of the matched firm. The free entry condition states that the flow cost of vacancy equals the expected profit of the firm in the future.

3.3.2 Matched firms

A firm-worker pair jointly produces a flow value of output $\phi(h,y;\omega)$. The firm pays the worker a wage $w(h,y;\omega)$. Then in the next period, conditional on the worker not caught committing a crime and not exogenously separated, the firm receives a continuation value $\tilde{\Pi}(h,y;\omega)$. Hence, the value of a matched firm is given by

$$\Pi(h,y;\omega)=\phi(h,y;\omega)-w(h,y;\omega)+\beta\left[1-\pi\lambda\left(1-H\left(\bar{m}^{W}(h,y;\omega)\right)\right)\right](1-s)\tilde{\Pi}(h,y;\omega) \tag{27}$$

where the continuation value is as follows

$$\tilde{\Pi}(h,y;\omega)=\begin{cases}\int_{h_{\min}}^{h_{\max}}\max\{\Pi(h',y;2),0\}\,dF^{\omega}(h'|h) & \text{when } \omega=2\\ \int_{h_{\min}}^{h_{\max}}\varphi_{PHI}^{W}\max\{\Pi(h',y;1),0\} & \\ +\left(1-\varphi_{PHI}^{W}\right)\max\{\Pi(h',y;0),0\}\,dF^{\omega}(h'|h) & \text{when } \omega\in\{0,1\}\end{cases} \tag{28}$$

3.4 Matching, surplus, and wage bargaining

Assume a standard constant-returns-to-scale matching function $M(u,v)=Au^{\alpha}v^{1-\alpha}$, then the meeting rates for unemployed workers and vacant firms are respectively,

$$\rho_y = \frac{M(u_y,v_y)}{u_y}=M(1,\theta_y)=A\theta_y^{1-\alpha} \tag{29}$$

$$q_y = \frac{M(u_y,v_y)}{v_y}=M\left(\theta_y^{-1},1\right)=A\theta_y^{-\alpha} \tag{30}$$

where $\theta_y=\frac{v_y}{u_y}$ is the market tightness of the market y. The total surplus from matching is the difference between the total matched values and the value of non-employment, which is the opportunity cost of employment:

$$\begin{aligned} S(h,y;\omega) &= V^{W}(h,y;\omega)+\Pi(h,y;\omega)-V^{NE}(h,y;\omega) \\ &= -\tau^{i}w(h,y;\omega)-\tau^{c}+\lambda K^{W}(h,y;\omega)+(1-\lambda)\beta\tilde{V}^{W}(h,y;\omega) \\ &\quad +\phi(h,y;\omega)+\beta(1-s)\left[1-\pi\lambda\left(1-H\left(\bar{m}^{W}(h,y;\omega)\right)\right)\right]\tilde{\Pi}(h,y;\omega) \\ &\quad -V^{NE}(h,y;\omega) \end{aligned} \tag{31}$$

Wages are determined by Nash bargaining

$$w(h,y;\omega)=\arg\max\left(V^{W}(h,y;\omega)-V^{NE}(h,y;\omega)\right)^{\eta}\left(\Pi(h,y;\omega)\right)^{1-\eta} \tag{32}$$

where η is the share of the matched surplus earned by the worker. The solution of the bargaining problem is

$$V^{W}(h,y;\omega)-V^{NE}(h,y;\omega)=\eta S(h,y;\omega) \tag{33}$$

which implies the wage is given by

$$\begin{aligned} w(h,y;\omega) &= \frac{1-\eta}{1-(1-\eta)\tau^i}\left[\tau^c + V^{NE}(h,y;\omega) - \lambda K^W(h,y;\omega) - (1-\lambda)\beta\tilde{V}^W(h,y;\omega)\right] \\ &+ \frac{\eta}{1-(1-\eta)\tau^i}\left[\phi(h,y;\omega) + \beta(1-s)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h,y;\omega)\right)\right)\right]\tilde{\Pi}(h,y;\omega)\right] \end{aligned} \tag{34}$$

In Appendix D, we lay out more details about the equilibrium and flow equations to close the model, as well as various measurements of crime and labor market in the economy.

4 Quantitative analysis

In this section, we calibrate the model to the US economy. We then evaluate the impact of health on the labor market outcomes and criminal behaviors of individuals.

4.1 Specifications and calibration

Following standard specification in the literature, the production function is assumed to be a product of the aggregate productivity factor z, health status h, and ability y:

$$\phi(h,y;\omega) = z^\omega h^{\phi_h^\omega} y^{\phi_y^\omega} \tag{35}$$

where z^ω, ϕ_h^ω, and ϕ_y^ω are parameters that depend on insurance status. That the productivity of workers depends on the health insurance status is consistent with the finding in Dizioli and Pinheiro (2016) that health insurance tends to enhance productivity. Also, the job search cost function of the unemployed workers is given by

$$\sigma(h) = \frac{\sigma}{1+h} \tag{36}$$

where σ is a parameter. The distribution of the value of crime is assumed to be uniform in $[m_{\min}, m_{\max}]$, hence we have for any $m \in [m_{\min}, m_{\max}]$,

$$H(m) = \frac{m}{m_{\max} - m_{\min}} \tag{37}$$

On the other hand, the health status is discretized into five states (Poor, Fair, Good, Very Good, and Excellent), as in the data. Where applicable, we enumerate the health status from 1 to 5. We

Table 3: Selected target statistics in the economy

Statistics	Data	Model
Labor market		
Job finding rate	52.0%	51.1%
Aggregate unemployment rate	4.55%	4.67%
Participation rate	75.8%	80.1%
Crime		
Crime rate	3.19%	3.18%
Average amount stolen per person	0.0128	0.0128
Health insurance coverage		
Employer-sponsored	0.699	0.688
Public, employed	0.194	0.189
Public, unemployed	0.654	0.651
Public, not-in-labor-force	0.909	0.909

derive $F^{\omega}(h'|h)$ from the transition matrix of health status obtained from the Medical Expenditure Panel Survey (MEPS; Blewett et al., 2018). Table E.1 shows the transition matrix, conditional on the health insurance status.

There are 28 parameters in the quantitative model. For some parameters, we use values standard in the literature. Then we calibrate the rest to the US data. Table 3 shows some of the target statistics in the baseline calibration, which will be described in detail below.

First, one period of the model is taken to be one year. The discount factor β is taken to be 0.9615, which is consistent with an annual discount rate of 4%. The bargaining power of the worker, η, is assumed to be 0.5, which is standard in the literature (Petrongolo and Pissarides, 2001). We assume Hosios condition, which implies $\alpha = \eta = 0.5$. The flow utility in prison x is zero, following Engelhardt et al. (2008).

For the parameters related to crime, we calibrate them to match the US property crime statistics in 2015, mostly following the approach in Burdett et al. (2003) and Engelhardt et al. (2008). For the distribution of the criminal value, we set $m_{\min} = 0$. Then, we choose $m_{\max}$ to match the average amount stolen per person from property crimes, which is 0.01276 (in terms of labor productivity) from the Crime in the US data in the Uniform Crime Reporting Program (UCR) published by the FBI. Following Engelhardt et al. (2008), we choose λ to match the annual property crime rate of 3.19%, which we obtain from dividing the number of incidents by the number of civilian noninstitutional population in the United States. The release rate from prison δ is taken to match the average time of 21 months served in state prison before the first release for property crime (Kaeble, 2018). This will imply that $\delta = \frac{12}{21} = 0.5714$. Finally, we take the clearance rate of property crime from the UCR data, which is 0.194, to be the value of π.

Labor market parameters are calibrated to match the US labor market statistics in 2015. Specifically, the parameter of matching efficiency A is chosen to match the annual job finding rate in the data, which is 52.0% in the Current Population Survey (CPS) microdata.[26] The exogenous separation rate s is taken to match the official unemployment rate (4.55%) in CPS. We also normalize the average labor productivity in the model to 1 by choosing the productivity factor z^{ω}. The unit of account in the calibrated model is in terms of labor productivity. In other words, one unit in the model corresponds to $\$121,746$ in 2015. The flow utility of non-employment b is set to be 0.4, following Shimer (2005). Moreover, the parameter σ in the job searching cost function is chosen to target the labor force participation rate in the CPS in 2015. Finally, the vacancy cost γ is taken to be 0.1, which is close to the figure of hiring cost in Hagedorn and Manovskii (2008).

We take the ability y to be consistent with the education groups in the data. Specifically, we specify five education groups (i.e., $J = 5$): Less than high school diploma (<HS), high school graduates (HS), some college (SC), college graduates (College), and more than college (>College). The initial joint distribution of education and health is taken from the PSID data and is shown in Table E.2. We calibrate $\{z^{\omega}, \phi_h^{\omega}, \phi_y^{\omega}\}$ to match some of the relative wages in the PSID data. We enumerate the education level from 1 to 5. Table E.3 shows the wage rate in the data and the model relative to that with the lowest education and health status. It turns out that the estimates of ϕ_h^{ω} and ϕ_y^{ω} are close to the wage elasticities of health and education in the data.[27] Lastly, the health insurance parameters $\left\{\varphi_{ESHI}, \varphi_{PHI}^{W}, \varphi_{PHI}^{U}, \varphi_{PHI}^{N}, \varphi_{PHI}^{P}\right\}$ are calibrated to jointly target the health insurance coverage rates in the PSID data.

Table 4 shows the values of parameters in the baseline calibration.

[26] The annual job finding rate is calculated from the fraction of unemployed workers who become employed 12 months later.

Obviously, the measure of annual job finding probability suffers from time aggregation error. Alternatively, one could calibrate the model at the monthly frequency and match other annual moments. It can be shown that the latter strategy yields similar quantitative results.

[27] By regressing log wage rate on log health and log education after enumerating the health status and education level from 1 to 5, we get that the wage elasticities are 0.25 and 0.60 respectively.

Table 4: Values of parameters (Baseline calibration)

Parameters	Meaning	Value	Target/source
β	Discount factor	0.9615	Annual $r = 4\%$
η	Bargaining power of the worker	0.5	Hosios condition
α	Matching elasticity	0.5	Petrongolo and Pissarides (2001)
x	Flow utility in prison	0	Engelhardt et al. (2008)
$m_{\min}$	Minimum of criminal value	0	Normalization
$m_{\max}$	Maximum of criminal value	0.7008	Average amount stolen per person (UCR)
λ	Probability of criminal opportunity	0.4458	Annual property crime rate (UCR/CPS)
δ	Release rate from the prison	0.5714	Average time served in state prison (UCR)
π	Apprehension rate	0.1940	Clearance rate of property crime (UCR)
A	Matching efficiency parameter	0.3302	Job finding rate (CPS)
s	Exogenous separation rate	0	Unemployment rate (CPS)
γ	Cost of maintaining vacancy	0.1	Hagedorn and Manovskii (2008)
b	Flow utility of non-employment	0.4	Shimer (2005)
σ	Parameter in the searching cost function	2.8	Labor force participation rate (CPS)
φ_{ESHI}	Rate of getting ESHI	0.1895	ESHI coverage (PSID)
φ_{PHI}^{W}	Rate of getting PHI, employed	0.6129	PHI coverage, employed (PSID)
φ_{PHI}^{U}	Rate of getting PHI, unemployed	0.4886	PHI coverage, unemployed (PSID)
φ_{PHI}^{N}	Rate of getting PHI, not-in-labor-force	0.9585	PHI coverage, not-in-labor-force (PSID)
φ_{PHI}^{P}	Rate of getting PHI, release from prison	0.5423	PHI coverage, release from prison (PSID)
z^{ω}	Productivity factor	$[0.1624, 0.3840, 0.3753]$	Relative wages (PSID)
ϕ_h^{ω}	Production elasticity of health status	$[0, 0.2854, 0.2299]$	Relative wages (PSID)
ϕ_y^{ω}	Production elasticity of education	$[0.4980, 0.5386, 0.6756]$	Relative wages (PSID)

4.2 Effects of health status on labor market outcomes

Figure 3 shows the joint distribution of health and employment statuses. We can see that most of the population in the economy possess no worse than a Good health status. Also, in general, conditional on health status, workers with better health are more likely to be employed and to participate in the labor market. In the model, workers with Poor health status never participate in the labor market. They would wait until their health improves.

Figure 4 shows several labor market statistics in the economy by health and health insurance status. For example, we calculate the labor market participation rate for each health and health insurance status. The result is shown in the top right panel. Not surprisingly, the participation rate is increasing with health status, with and without health insurance. Also, the participation rate is slightly lower for those with health insurance compared to those uninsured.[28] The employment-population ratio, shown in the bottom left panel, shows a similar story.

What is perhaps surprising is the average value of workers by employment status, shown in the bottom right panel. The average values are computed by taking the average of the respective values

[28] This is because of the compositional effect: not-in-labor-force workers have a higher chance of getting public health insurance than the unemployed ones, and so the insured group is more likely to contain individuals who are out of labor force.

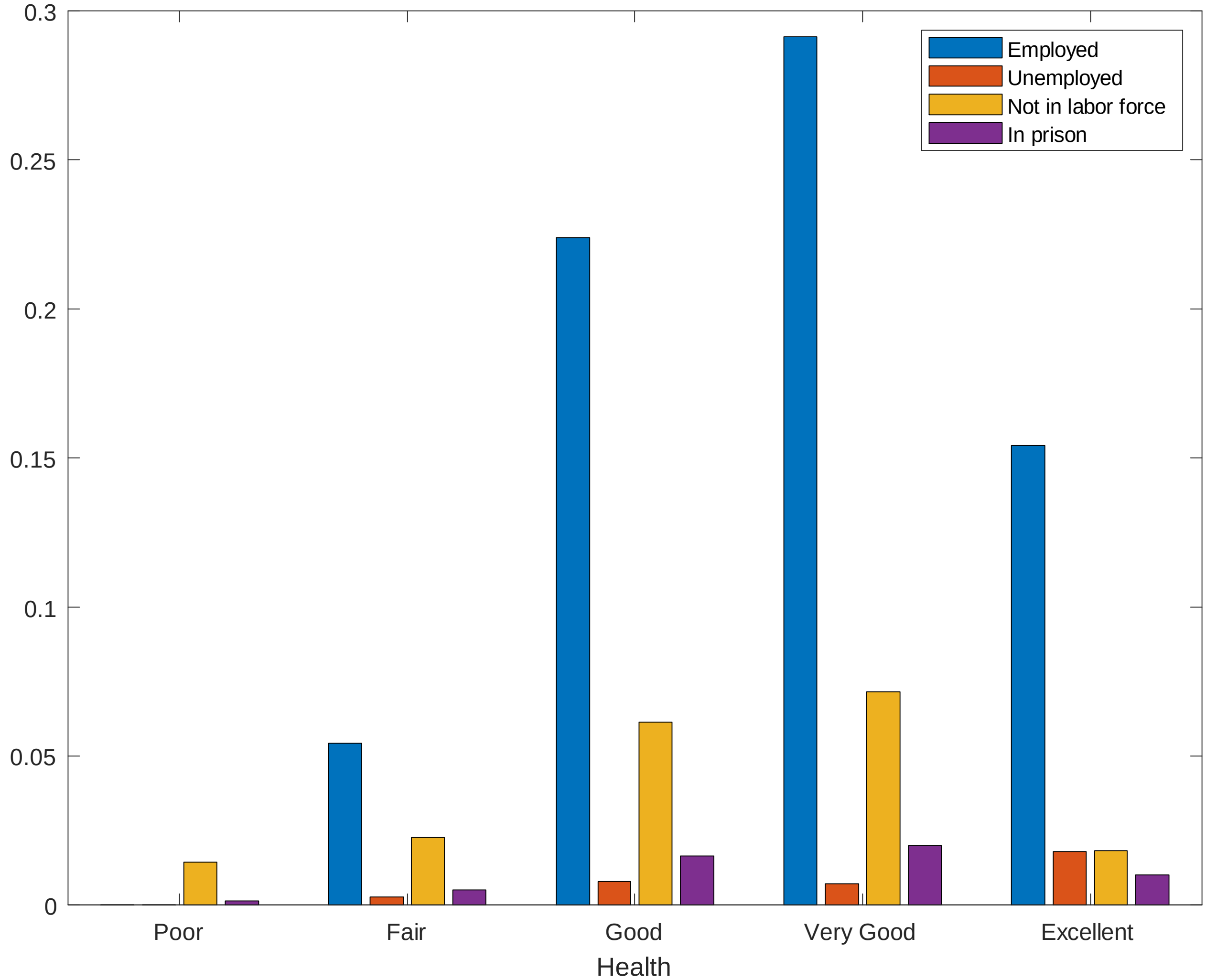

Figure 3: Joint distribution of health and employment statuses

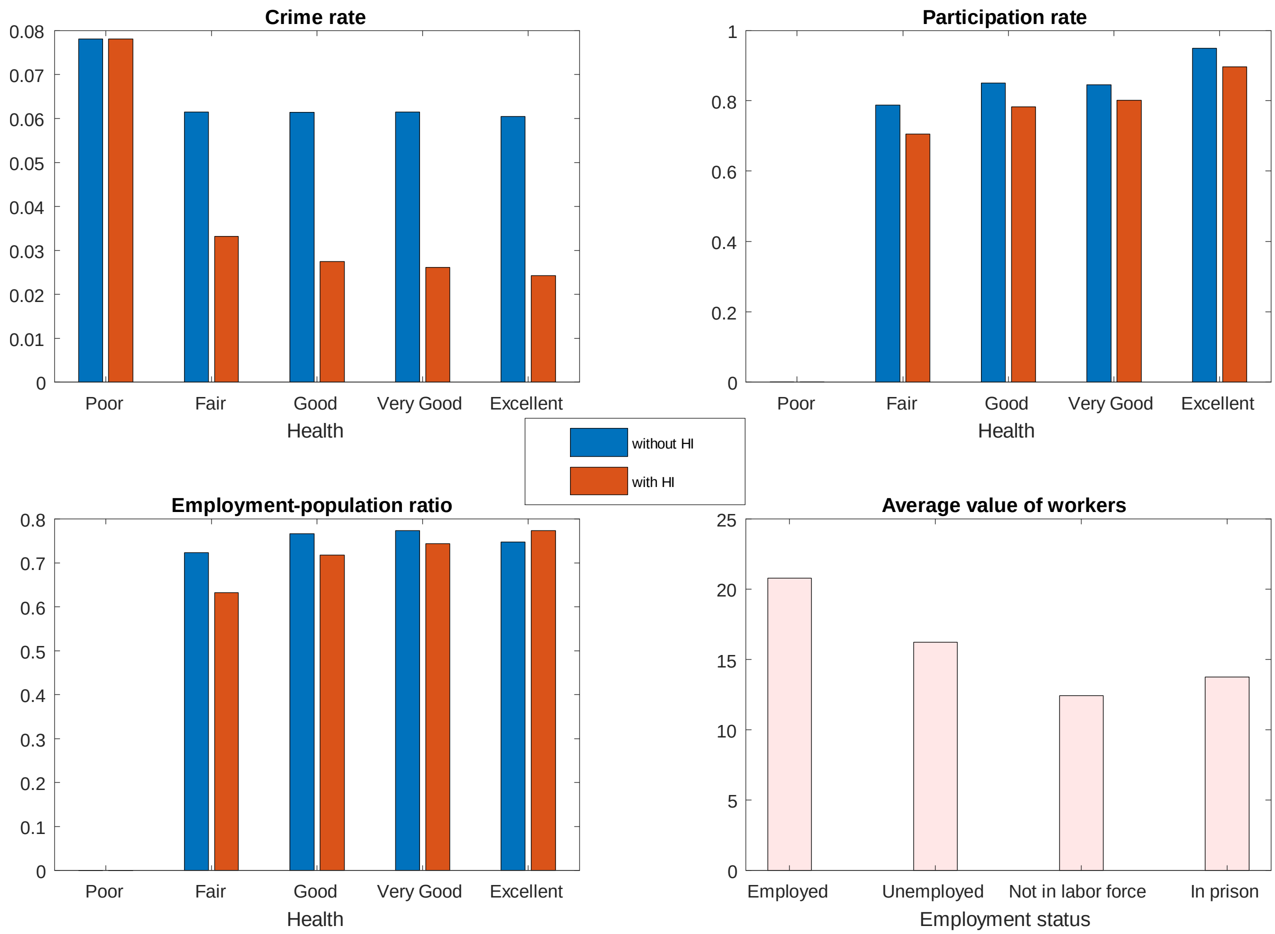

Figure 4: Baseline results

given by (5), (10), (13), and (13), weighted by the corresponding density of workers by health and ability. As long as individuals are out of prison, the mean value for the employed is the highest, and that for those not-in-labor-force is the lowest. However, the mean value for imprisoned individuals is slightly higher than those out of the labor force. This is due to the selection effect: the health and productivity of individuals in the not-in-labor-force state are low since they are a group of people who are not actively looking for a job. Hence, the average productivity in the non-in-labor-force state is lower than that of the in-prison state.

In summary, healthier workers have higher labor force participation, employment-population ratio,[29] and productivity. These facts are consistent with our empirical analysis and those in the literature.

4.3 Effects of health status on criminal behaviors

The effects of health on individuals' criminal behavior can be observed in Figure 5, which shows the crime rate by health status and education level in the model. The left panel shows the crime rate of those who are uninsured, and the right panel shows that of those who have health insurance. In general, the crime rate is decreasing with both better health status and higher education level. This is because of the relative higher opportunity cost of committing a crime associated with better health and higher education level, as explained in Section 3. Note that health insurance has a huge impact: the crime rate of those with health insurance tend to be much more sensitive to health status and education level. This is because the production elasticities of health and education are higher for those with health insurance. Hence, the opportunity cost of committing a crime would, in turn, be more dependent on health and education. The crime rate of those with health insurance, Excellent health status, and more than college education is only about one-sixth of those groups with the highest crime rates. The top left panel of Figure 4 shows a very similar picture. We conclude that criminal activity is negatively related to the health status and education level of the worker quantitatively, and is more so for those with health insurance.

[29] Interestingly, while not shown in the figure, the unemployment rate by health status shows no clear pattern. This is likely due to the participating decision of uninsured individuals. While better health status improves the unemployment rate among insured participants, due to higher job finding rate and lower job separation rate, it would also induce more uninsured individuals to participate, which drives up the unemployment rate.

Figure 5:
Crime rate by health, education, and health insurance status

5 Policy analysis

The model developed in the previous sections provides us with a perfect laboratory for public policy analysis. We now perform a wide range of policy experiments in the baseline model and evaluate their impacts on the labor market outcomes and criminal behaviors of workers with different health status. This section investigates the effects of health insurance policies, unemployment benefits, and criminal sentencing policy. In Appendix F, we show the results of additional policy analyses. The baseline results are shown in Figure 4. In the subsequent figures, we show the changes from the baseline levels following each policy experiment.

5.1 Health insurance coverage

Figure 6 shows the impact of health insurance policies on the labor market and crime variables. We consider two policy experiments. The first policy, which we refer to as Medicare-for-all, provides everyone with public health insurance (i.e. $\varphi^{W}_{PHI} = \varphi^{U}_{PHI} = \varphi^{N}_{PHI} = 1$). The second policy is

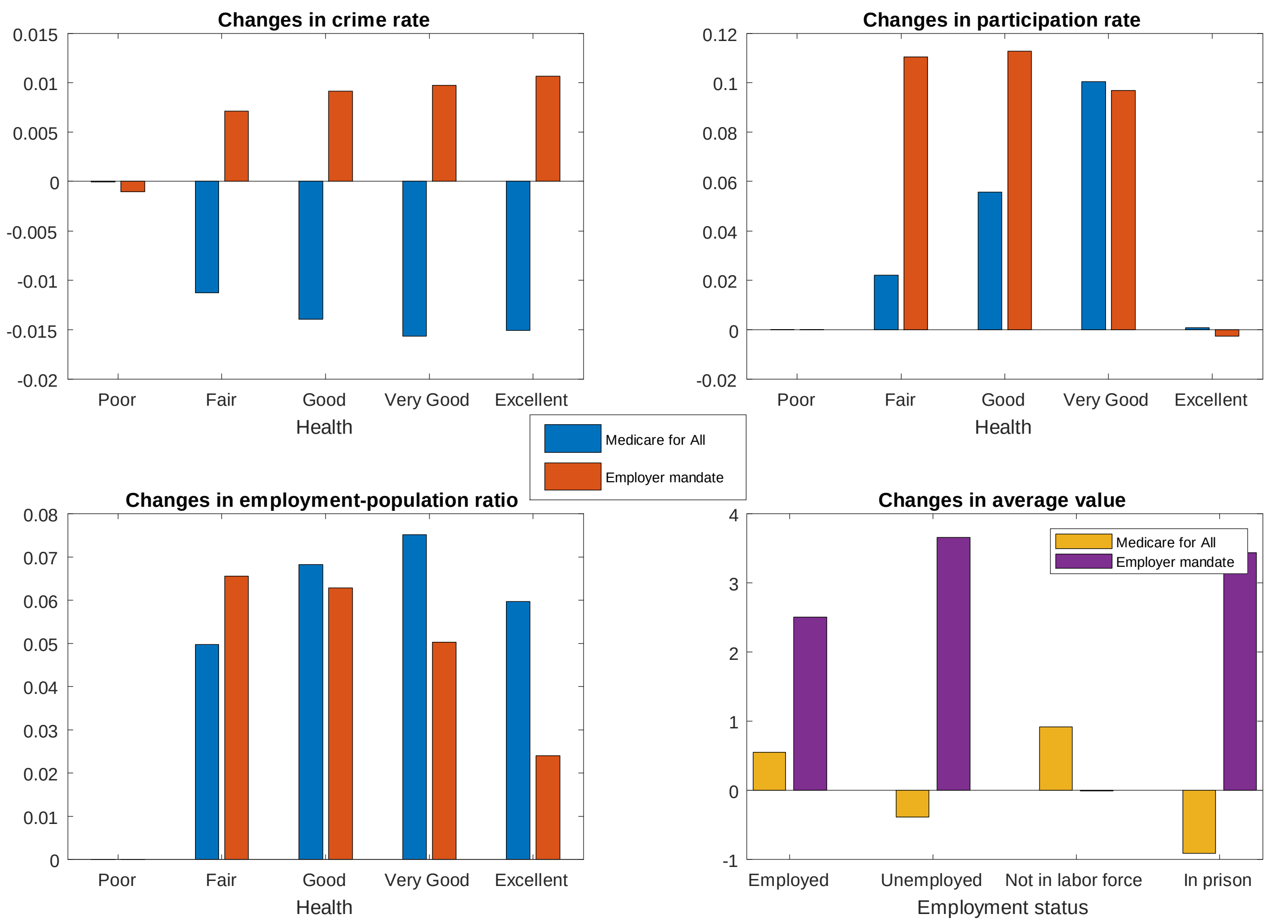

Figure 6: Effects of health insurance coverage

the Employer Mandate under the ACA, where now all jobs provide employer-sponsored health insurance (i.e. $\varphi^{W}_{ESHI} = 1$).

First, Medicare-for-all effectively lowers the crime rate of workers with every health status. Intuitively, with higher health insurance coverage, the expected health status of workers becomes better. This implies a higher job-finding rate and a higher wage rate, essentially increasing the opportunity cost of committing a crime. Consequently, there is less incentive to commit a crime. Also, the reduction of the crime rate is increasing with health. This is because healthier workers, who are more likely to participate in the labor market, benefit more in the job market from improving health, thereby increasing the opportunity cost of committing a crime more. The result that higher health insurance coverage lowers the crime rate is consistent with Vogler (2020) and He and Barkowski (2020), who have shown that the ACA contributes to the lower crime rate.

Also, health insurance has a positive impact on the participation rate. For example, if universal health insurance is to be implemented, the participation rate of a worker with a Very Good health status would increase by about 10 percentage points, which is quantitatively significant. This is consistent with the empirical finding that as the workers become healthier, they are more likely to participate in the job market. The effect on workers with Excellent health status is negligible since their participation rate is already high under the baseline case. Similar results are found in the employment-population ratio for those with health statuses Fair, Good, or Very Good. However, there is a significantly positive effect on individuals with Excellent health status due to the lower endogenous job separation rate.

Finally, Medicare-for-all increases the average values of those out of prison and decreases the value of prisoners, primarily because now everyone is insured, either by public health insurance or by employer-sponsored health insurance.

The Employer Mandate affects outcomes differently. Under the policy, the crime rate increases, and its magnitude is larger for individuals with better health status. The reason is clear from the changes in average values. The Employer Mandate increases the values of the employed, the unemployed, and prisoners, but the increase is larger for the unemployed and prisoners. This is because a large part of employees is already insured by ESHI under the baseline, the benefit of the new policy is limited for the employed. As the unemployed and prisoners can get health insurance when they become employed in the future, the average value increases through the expected future

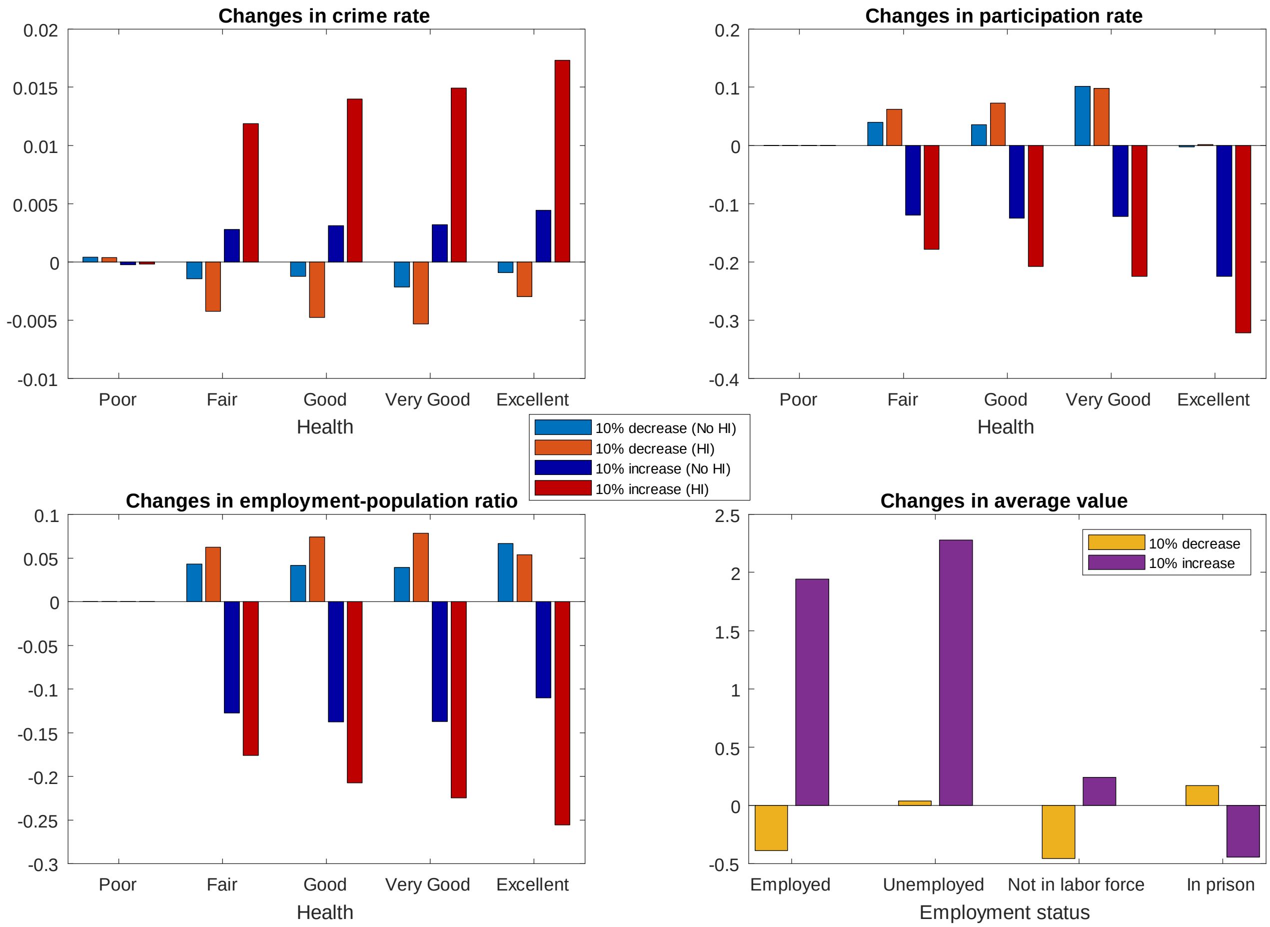

Figure 7: Effects of changing unemployment benefit

values. Hence, the value of prisoners increases more than the employed, which increases the crime rate of the employed workers. As a result, the aggregate crime rate increases for each health status since most of the labor force is employed.

5.2 Unemployment benefits

The government can increase the flow utility of non-employment, for example, by changing the unemployment benefit. In the baseline calibration, we set $b = 0.4$. Here we consider the cases when the flow utility is 10% higher (i.e. $b = 0.44$) and lower (i.e. $b = 0.36$) and study the impacts on the labor market and criminal behaviors.

Figure 7 shows the impact of different flow utility of non-employment on the criminal and labor market outcomes.[30] When the flow utility increases, the value of unemployed workers increases the

[30] There is no change of labor market outcomes for the Poor health people since they do not participate in the

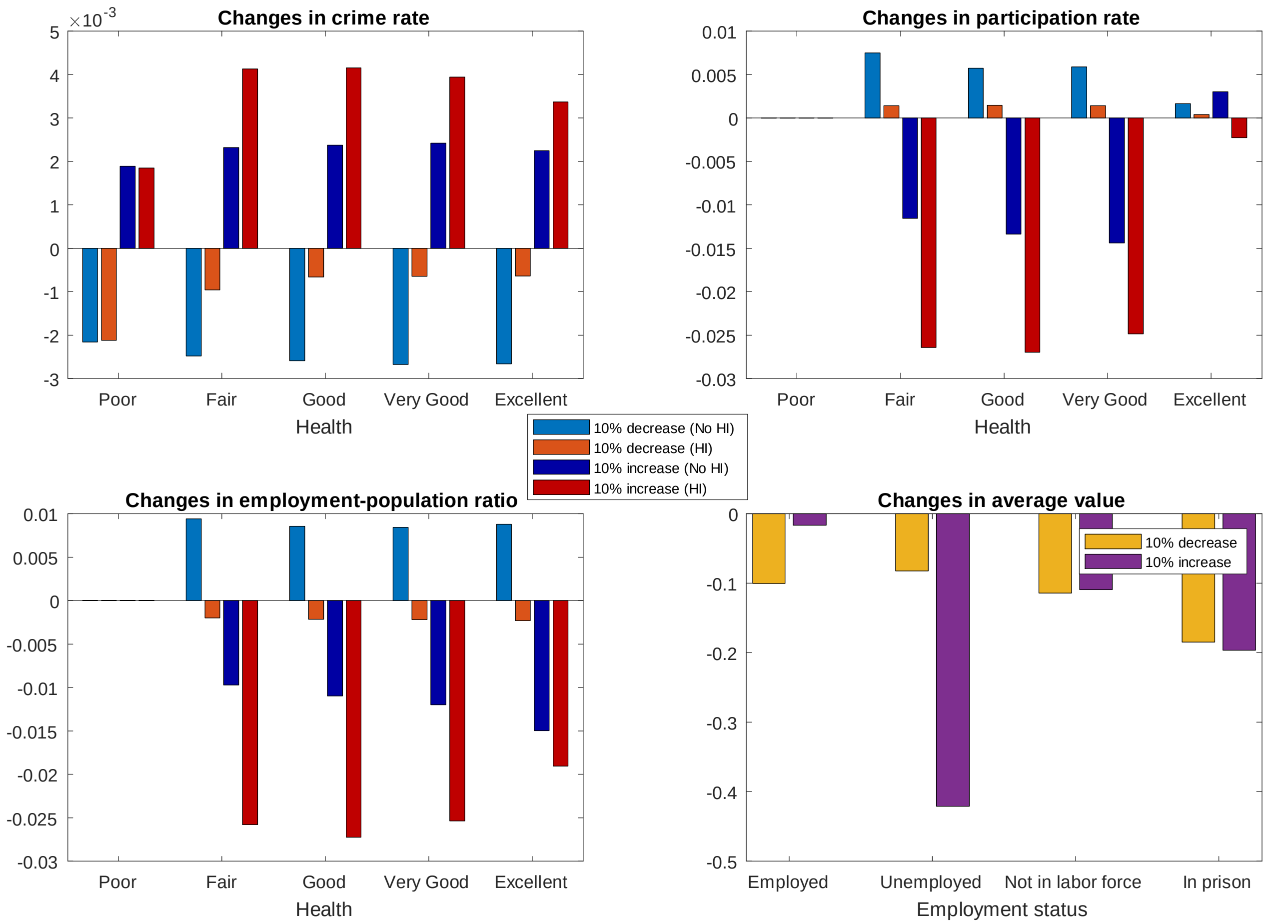

Figure 8: Effects of changing release rate from prison

most, followed by employed workers through the wage bargaining process. Accordingly, employment becomes less attractive for individuals, and they have less incentive to be employed. Therefore, we observe a decrease in the participation rate and the employment-population ratio.

After a 10% increase in the flow utility, Poor health people are less likely to commit a crime for fear of losing the increased flow utility by incarceration. Hence, their crime rate slightly decreases. However, for relatively healthy categories, the crime rate increases through lower participation in the labor market. In general, changes in the flow utility affect the insured individuals more than the uninsured.

5.3 Sentencing policy

One of the most important crime related policies is the sentencing policy. For instance, in the United States the Federal Sentencing Guidelines set out a uniform sentencing policy for serious felonies and misdemeanors. In our model, it can be captured by a change in the release rate δ. In the baseline case, we have $\delta = 0.571$, which corresponds to an average of approximately 21 months in prison. Figure 8 shows the impact of different release rates on the labor market and criminal outcomes. We consider two cases: when δ is 10% higher ($\delta = 0.629$) or 10% lower ($\delta = 0.514$), which, compared to the baseline, corresponds to about two months shorter and longer average sentence, respectively.

First, when the sentencing policy becomes harsher (i.e., a lower release rate), it is clear that the cutoff level of crime increases, and fewer crimes are committed. The opposite is true when the release rate increases.

The release rate also changes labor market outcomes. With more lenient sentencing, both the participation rate and the employment-population ratio decrease. Note that a sentencing policy affects imprisoned individuals and individuals from other states indirectly through the expected values. A more lenient sentencing policy leads to individuals relying less on employment than criminal activities, which decreases the average value of each state. With harsher sentencing, the changes in average values are mainly driven by compositional changes: it reduces the incentive for the marginal individuals, who have a below-average productivity in each of the out of prison state, to commit a crime. As a result, all states experience a slight decrease in the average value.[31]

5.4 Effects on aggregate variables

Table 5 shows the effects of the public policies on the aggregate statistics, including the unemployment rate (u), crime rate (cr), labor force participation rate (pr), aggregate output (GDP),[32]

labor market both in the baseline and under unemployment benefit policy.

[31] For example, a marginal unemployed individual who is deciding to commit a crime has a below average productivity within the unemployment pool but above average productivity among the prisoners. Therefore, discouraging the marginal individual to commit a crime reduces the average productivity of both the unemployed and prisoners states.

[32] Since we have normalized the population to unity, our notion of GDP and GDP per capita are equivalent.

Table 5: Policy impact on aggregate variables

		Policy								
(a) Level	Baseline	Unemp. benefits		HI coverage		Income tax	Release rate		Apprehension rate	
		$b \downarrow 10\%$	$b \uparrow 10\%$	Medicare-for-all	Emp. mandate	$\tau^i = 10\%$	$\delta \downarrow 10\%$	$\delta \uparrow 10\%$	$\pi \downarrow 10\%$	$\pi \uparrow 10\%$
u	4.67%	4.29%	4.11%	5.31%	6.12%	4.21%	4.59%	5.31%	5.73%	4.73%
cr	3.18%	2.78%	4.47%	1.74%	4.12%	4.52%	3.09%	3.58%	3.59%	3.35%
pr	80.13%	86.62%	57.85%	86.00%	88.28%	57.56%	80.31%	78.24%	76.62%	80.18%
GDP	0.72	0.77	0.57	0.82	0.84	0.56	0.72	0.70	0.68	0.72
Gini	0.26	0.26	0.28	0.21	0.26	0.30	0.26	0.27	0.28	0.26
Welfare	21.32	21.72	20.07	23.08	22.76	18.25	21.20	20.95	20.33	21.37
(b1) % point change										
u		-0.4	-0.6	0.6	1.4	-0.5	-0.1	0.6	1.1	0.1
cr		-0.4	1.3	-1.4	0.9	1.3	-0.1	0.4	0.4	0.2
pr		6.5	-22.3	5.9	8.1	-22.6	0.2	-1.9	-3.5	0.1
(b2) % change										
GDP		6.7%	-21.4%	13.6%	16.4%	-21.9%	-0.5%	-2.8%	-6.5%	0.2%
Gini		-2.7%	7.9%	-21.6%	0.1%	16.2%	1.0%	3.1%	7.2%	0.5%
Welfare		1.9%	-5.9%	8.3%	6.7%	-14.4%	-0.5%	-1.8%	-4.7%	0.2%

Gini coefficient,[33] and welfare.[34] Panel (a) shows the level of each statistic, whereas panel (b1) and panel (b2) show respectively the percentage point change and percentage change for each measure.

A 10% increase in unemployment benefit, for example, would lower rates of unemployment and labor force participation but increase the crime rate in the economy, though the effects are arguably small in magnitude. As a result of the lower participation, there will also be a drop in the aggregate output.[35] The Gini coefficient indicates that the policy widens the inequality by 7.9% and lowers overall welfare by 5.9%.

On the aggregate level, both Medicare-for-all and Employer Mandate policies would increase the participation rate. However, while Medicare-for-all decreases crime rates, the effects of the Employer Mandate are the opposite. The increase in participation and a higher production elasticity for insured workers imply that the aggregate output would be higher. Our model estimates the crime rate would decrease by 1.4 percentage points, and the GDP would increase by 13.6% when the Medicare-for-all policy is introduced.[36] Medicare-for-all decreases inequality by 21.6% by increasing individuals' productivity. The Employer Mandate policy increases the GDP by a comparable amount. However, the effect on inequality is minimal since the productivity increase is limited to the employed.

The introduction of income tax, in general, has a quantitatively significant impact on the economy.[37] For instance, a 10% income tax would reduce the number of unemployed workers in the economy by half a percentage point and increase criminal activities by more than one percentage point. It would also lower participation and hence aggregate output by about 20%.

Lastly, crime policies have moderate impacts on labor market outcomes and aggregate output. For example, a 10% increase in the release rate would increase the unemployment rate by 0.6 percentage point and crime rate by 0.4 percentage point, and it decreases the participation rate by

[33] We compute the Gini coefficients using the flow income values of all individuals in the economy. Specifically, we take b and x as the income of the non-employed and the imprisoned individuals respectively. The Gini coefficients computed using only the employed have similar values.

[34] The welfare function is defined as the sum of individual values and the firm's value weighted by their measures. This welfare function is consistent with an alternative welfare measure of net product in the economy and gives qualitatively similar results.

[35] Interestingly, there is also a decrease in the unemployment rate. This is due to the lower participation of workers with relatively bad health status.

[36] Using different model specifications, Aizawa (2019) and See (2019) find a much smaller or even negative productivity effect of ACA and Medicare-for-all policies. In our model, the relatively large effect on the aggregate output is due to a combination of individual productivity gains (about 4%) and the increased participation in the labor market.

[37] Here we introduce an income tax only on the wages of the employed workers. The tax revenue is not used for any welfare-improving activity. See Appendix F for the detailed analysis.

1.9 percentage points. It would also lower the aggregate output by 2.8%.

6 Conclusion

In this paper, we study the links between health, crime, and the labor market. Our estimations show that healthier workers, on average, fare better labor market outcomes and engage in fewer criminal activities. We build a model to show the economic mechanism of these relationships: other things equal, the better job prospects of healthier workers raise the opportunity cost of criminal activities, thereby reducing the incentive of committing crimes. We also perform a wide range of policy experiments. For instance, while both Medicare-for-all and Employer Mandate under the ACA increase economic output, they have different and important side effects on crime and inequality. This shows that a unified model of health, crime, and the labor market is essential for public policy analysis. We believe the framework developed in this paper can be useful for further public policy analysis.

A few promising extensions of our model are in order. For instance, our model abstracts from the life-cycle and behavioral aspects of health dynamics. Theoretically, health can be treated as an investment good as in the classical Grossman (1972) model. We do not include this factor because the relationship between medical spending and health is not conclusive in the empirical literature. Modeling health endogenously by allowing, for example, individuals choosing their medical spending would certainly help understand the dynamics of health. Also, while the health insurance coverage is modeled stochastically by exogenous probability, one can extend the model to include health insurance decisions.[38] Understanding the life-cycle dynamics of criminal behaviors by modeling the health dynamics more realistically is a worthwhile undertaking for future investigation.

References

Aalsma, M. C., White, L. M., Lau, K. S., Perkins, A., Monahan, P., and Grisso, T. (2015). Behavioral health care needs, detention-based care, and criminal recidivism at community reentry

[38] Dey and Flinn (2005), for example, explicitly include health insurance choice in their model to study the impact of health insurance on the labor market outcomes.

from Juvenile detention: A multisite survival curve analysis. *American Journal of Public Health*, 105(7):1372–1378.

Aizawa, N. (2019). Labor market sorting and health insurance system design. *Quantitative Economics*, 10(4):1401–1451.

Aizawa, N. and Fang, H. (2020). Equilibrium Labor Market Search and Health Insurance Reform. *Journal of Political Economy*, 128(11):4258–4336.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2):169–217.

Blewett, L. A., Drew, J. A. R., Griffin, R., Williams, K. C., and Backman, D. (2018). IPUMS Health Surveys: Medical Expenditure Panel Survey, Version 1.0 [dataset].

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Burdett, K., Lagos, R., and Wright, R. (2003). Crime, Inequality, and Unemployment. *American Economic Review*, 93(5):1764–1777.

Burdett, K., Lagos, R., and Wright, R. (2004). An on-the-job search model of crime, inequality, and unemployment. *International Economic Review*, 45(3):681–706.

Cai, L. (2009). Effects of health on wages of Australian men. *Economic Record*, 85(270):290–306.

Cuellar, A. E., Markowitz, S., and Libby, A. M. (2004). Mental health and substance abuse treatment and juvenile crime. *The Journal of Mental Health Policy and Economics*, 7(2):59–68.

Cuellar, A. E., McReynolds, L. S., and Wasserman, G. A. (2006). A cure for crime: Can mental health treatment diversion reduce crime among youth? *Journal of Policy Analysis and Management*, 25(1):197–214.

Currie, J. and Madrian, B. C. (1999). Health, health insurance and the labor market. In *Handbook of Labor Economics*, volume 3, chapter 50, pages 3309–3416.

Dey, M. S. and Flinn, C. J. (2005). An Equilibrium Model of Health Insurance Provision and Wage Determination. *Econometrica*, 73(2):571–627.

Dizioli, A. and Pinheiro, R. (2016). Health insurance as a productive factor. *Labour Economics*, 40:1–24.

Edmark, K. (2005). Unemployment and Crime: Is There a Connection? *Scandinavian Journal of Economics*, 107(2):353–373.

Engelhardt, B., Rocheteau, G., and Rupert, P. (2008). Crime and the labor market: A search model with optimal contracts. *Journal of Public Economics*, 92(10-11):1876–1891.

Fazel, S. and Baillargeon, J. (2011). The health of prisoners. *The Lancet*, 377(9769):956–965.

Fazel, S. and Grann, M. (2006). The Population Impact of Severe Mental Illness on Violent Crime. *American Journal of Psychiatry*, 163(8):1397–1403.

Fazel, S., Gulati, G., Linsell, L., Geddes, J. R., and Grann, M. (2009). Schizophrenia and Violence: Systematic Review and Meta-Analysis. *PLoS Medicine*, 6(8):e1000120.

Frank, R. and McGuire, T. (2010). Mental Health Treatment and Criminal Justice Outcomes. *NBER Working Paper*.

García-Gómez, P. (2011). Institutions, health shocks and labour market outcomes across Europe. *Journal of Health Economics*, 30(1):200–213.

García-Gómez, P., Jones, A. M., and Rice, N. (2010). Health effects on labour market exits and entries. *Labour Economics*, 17(1):62–76.

García-Gómez, P., van Kippersluis, H., O'Donnell, O., and van Doorslaer, E. (2013). Long-Term and Spillover Effects of Health Shocks on Employment and Income. *Journal of Human Resources*, 48(4):873–909.

Garibaldi, P. and Wasmer, E. (2005). Equilibrium Search Unemployment, Endogenous Participation, and Labor Market Flows. *Journal of the European Economic Association*, 3(4):851–882.

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

Hagedorn, M. and Manovskii, I. (2008). The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited. *American Economic Review*, 98(4):1692–1706.

Halla, M. and Zweimüller, M. (2013). The effect of health on earnings: Quasi-experimental evidence from commuting accidents. *Labour Economics*, 24(0):23–38.

He, Q. and Barkowski, S. (2020). The effect of health insurance on crime: Evidence from the Affordable Care Act Medicaid expansion. *Health Economics*, 29(3):261–277.

Hosseini, R., Kopecky, K., and Fed Kai Zhao, A. (2018). The Evolution of Health over the Life Cycle. *Working paper*.

Jäckle, R. and Himmler, O. (2010). Health and Wages: Panel Data Estimates Considering Selection and Endogeneity. *Journal of Human Resources*, 45(2):364–406.

Kaeble, D. (2018). Time Served in State Prison, 2016. Technical Report November, Bureau of Justice Statistics.

Kinner, S. A. and Wang, E. A. (2014). The Case for Improving the Health of Ex-Prisoners. *American Journal of Public Health*, 104(8):1352–1355.

Kitazawa, Y. (2012). Hyperbolic Transformation and Average Elasticity in the Framework of the Fixed Effects Logit Model. *Theoretical Economics Letters*, 02(02):192–199.

Lin, M.-J. (2008). Does Unemployment Increase Crime? *Journal of Human Resources*, 43(2):413–436.

Marcotte, D. E. and Markowitz, S. (2010). A cure for crime? Psycho-pharmaceuticals and crime trends. *Journal of Policy Analysis and Management*, 30(1):29–56.

Mortensen, D. T. and Pissarides, C. A. (1994). Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies*, 61(3):397–415.

Nakajima, M. and Tuzemen, D. (2017). Health-Care Reform or Labor Market Reform? A Quantitative Analysis of the Affordable Care Act. *SSRN Electronic Journal*, pages 1–64.

Otsu, Y. (2016). Crime, Search, and Ex-offenders' Accessibility to Labor Markets. *LABOUR*, 30(4):393–414.

Pashchenko, S. and Porapakkarm, P. (2013). Quantitative analysis of health insurance reform: Separating regulation from redistribution. *Review of Economic Dynamics*, 16(3):383–404.

Petrongolo, B. and Pissarides, C. A. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature*, 39(2):390–431.

Regenstein, M. and Rosenbaum, S. (2014). What The Affordable Care Act Means For People With Jail Stays. *Health Affairs*, 33(3):448–454.

Roodman, D. (2009). How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal: Promoting communications on statistics and Stata*, 9(1):86–136.

Schnepel, K. T. (2018). Good Jobs and Recidivism. *The Economic Journal*, 128(608):447–469.

Schroeder, R. D., Hill, T. D., Haynes, S. H., and Bradley, C. (2011). Physical health and crime among low-income urban women: An application of general strain theory. *Journal of Criminal Justice*, 39(1):21–29.

See, K. (2019). "Medicare for All": The Macroeconomic and Welfare Consequences of Expanding Public Health Insurance. *Working paper.*

Shimer, R. (2005). The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *American Economic Review*, 95(1):25–49.

Stephens, M. and Toohey, D. (2018). The Impact of Health on Labor Market Outcomes: Experimental Evidence from MRFIT. *NBER Working Paper.*

Trevisan, E. and Zantomio, F. (2016). The impact of acute health shocks on the labour supply of older workers: Evidence from sixteen European countries. *Labour Economics*, 43(27):171–185.

Vogler, J. (2020). Access to Healthcare and Criminal Behavior: Evidence from the ACA Medicaid Expansions. *Journal of Policy Analysis and Management*, 39(4):1166–1213.

Wen, H., Hockenberry, J. M., and Cummings, J. R. (2014). The Effect of Substance Use Disorder Treatment Use on Crime: Evidence from Public Insurance Expansions and Health Insurance Parity Mandates. *NBER Working Paper*.

Wen, H., Hockenberry, J. M., and Cummings, J. R. (2017). The effect of Medicaid expansion on crime reduction: Evidence from HIFA-waiver expansions. *Journal of Public Economics*, 154:67–94.

Wilper, A. P., Woolhandler, S., Boyd, J. W., Lasser, K. E., McCormick, D., Bor, D. H., and Himmelstein, D. U. (2009). The health and health care of US prisoners: Results of a nationwide survey. *American Journal of Public Health*, 99(4):666–672.

Yang, C. S. (2017). Local labor markets and criminal recidivism. *Journal of Public Economics*, 147:16–29.

Appendix (for online publication)

A Additional empirical results

This appendix includes the detailed data cleaning process and additional robustness checks for the empirical results. In addition, we show the empirical relationship between health and labor market performance.

A.1 Crime regressions

A.1.1 Data selection

The sample selection process for the NLSY97 data is shown in Table A.1. The total number of observations in NLSY97 is originally 134,760. We drop those non-interview observations. Also, observations with missing education, health, marital status, employment status, and drug usage[39] are dropped. We then keep those individuals with age at or above 16. Finally, since we use forward crime variables in our regression for more consistent timing, we drop those observations when the forward variables are not available. Table A.2 shows the summary statistics of the sample.

Table A.1: Sample selection (NLSY 97)

Reason of dropping sample	Dropped	Total
Original		134760
Non-interview	17802	116958
Missing (Education)	1197	116958
Missing (Health)	48	115713
Missing (Marital status)	130	115583
Missing (Cocaine)	19141	96442
Age <16	1619	94823
Missing (forward property crime)		
Missing (forward violent behavior)	12047	82776
Summary Statistics		82776

A.1.2 Robustness checks

To check the relationship between health and crime further, Table A.3 and A.4 show the regression of property crime and violent behaviors with additional controls. The first column repeats the

[39] The relevant question is "Excluding marijuana and alcohol, since the date of the last interview, have you used any drugs like cocaine or crack or heroin, or any other substance not prescribed by a doctor, in order to get high or to achieve an altered state?"

Table A.2: Summary statistics (NLSY97)

N= 82776	Mean	SD	Min	Max	Poor	Fair	Good	Very good	Excellent
Age	22.6	3.64	16	30	24.1	23.3	22.9	22.7	22.1
Female	0.49	0.50	0	1	0.68	0.55	0.54	0.50	0.42
Black	0.15	0.36	0	1	0.17	0.21	0.16	0.13	0.15
Hispanic	0.13	0.33	0	1	0.14	0.15	0.14	0.12	0.12
Mixed race	0.013	0.11	0	1	0.026	0.019	0.013	0.011	0.014
Marriage	0.19	0.39	0	1	0.19	0.18	0.19	0.20	0.17
Years of education	11.4	1.13	0	12	10.8	11.0	11.3	11.5	11.5
Health	3.88	0.93	1	5	1	2	3	4	5
Log wage rate	2.06	1.08	-6.55	8.57	2.01	1.99	2.05	2.09	2.03
Violent behavior	0.027	0.16	0	1	0.076	0.040	0.031	0.021	0.026
Property crime	0.028	0.16	0	1	0.043	0.037	0.028	0.028	0.026
Hard drug usage	0.060	0.24	0	1	0.099	0.084	0.072	0.060	0.042

baseline regression result. The second column controls the employment status at the interview date, since health could affect criminal behavior through labor market outcomes. However, the estimated coefficient for the employment status is negative but statistically insignificant. The estimated coefficients for health measures still show a negative association. This result implies that employment status alone may not capture the labor market channel. The third column controls drug usage, since drug could affect both health and criminal behavior. Although the drug usage is positively associated with criminal behavior, the estimated result still confirms the association between better health and criminal behaviors. In the fourth column, we control both employment and drug usage, but the coefficients for health status remain statistically significant under the specification for both property crime and violent crime. The last two columns are to check the sensitivity to wages. The fifth column repeats the same specification in the first column with the sample with the wage variable, and in this subsample, we also confirm the relationship between better health and fewer crimes. The sixth column controls the log of wages but we still confirm the same relationship.

A.2 Health and labor market outcomes

This section documents the relationship between health and labor market outcomes using microdata from the Panel Study of Income Dynamics (PSID).

A.2.1 Data

The data from PSID is a nationally representative longitudinal sample of households and individuals in the United States. We check the relationship between health and labor market outcomes using the PSID dataset from 1999 to 2015. PSID contains self-reported health status for the head of households, which is reported as Excellent, Very Good, Good, Fair, or Poor health status. Since PSID only contains health information about the household heads, we focus on the head sample whose age is between 22 and 65. We exclude out-of-labor-force observations and use employed and unemployed observations for regression to focus on the labor market. Table A.5 summarizes the detailed data selection process, and summary statistics are in Table A.6. On average, healthier individuals acquire more education, are more likely to be employed, are less likely to be unemployed, earn higher wages, and work longer hours.

The sample selection process for the PSID data is shown in Table A.5. We start with 79,682 observations of household heads aged 22 to 65. We then drop those observations with missing health status, race, education, marital status, and employment status. This step leaves us 73,684 observations. Finally, for both the employment and wage regressions, we use the forward variable so that the timing is consistent. Additionally, for the wage regression, we trim the observations

Table A.3: Property crime regression on health status (NLSY97)

	(1)	(2)	(3)	(4)	(5)	(6)
Property crime	Logit	Logit	Logit	Logit	Logit	Logit
Fair	-0.654	-0.646	-0.663	-0.655	-1.184	-1.186
	(0.302)	(0.302)	(0.302)	(0.302)	(0.492)	(0.492)
Good	-0.845	-0.835	-0.847	-0.837	-1.401	-1.403
	(0.299)	(0.299)	(0.300)	(0.300)	(0.487)	(0.487)
Very Good	-0.869	-0.858	-0.866	-0.855	-1.420	-1.421
	(0.303)	(0.303)	(0.303)	(0.303)	(0.490)	(0.490)
Excellent	-0.738	-0.729	-0.728	-0.720	-1.355	-1.356
	(0.306)	(0.306)	(0.306)	(0.307)	(0.497)	(0.497)
Employment		-0.0653		-0.0667		
		(0.0662)		(0.0662)		
Hard Drug Use			0.212	0.212		
			(0.0837)	(0.0837)		
Log of wage						-0.00985
						(0.0422)
Control	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y	Y	Y
Observations	14,391	14,391	14,391	14,391	5,750	5,750
Number of individuals	1,409	1,409	1,409	1,409	767	767

Note: Robust standard errors in parentheses. The estimated coefficients are shown in the table. The dependent variable *property crime* is made from questions that ask if stealing money more/less than 50 dollars or if committing other property crime. Year FE for logit models are two-year dummies. Control variables are a female dummy, race dummies, age, age squared, years of education, and marital status.

Table A.4: Violent crime regression on health status (NLSY97)

Violent behavior	(1) Logit	(2) Logit	(3) Logit	(4) Logit	(5) Logit	(6) Logit
Fair	-0.544	-0.537	-0.544	-0.537	-0.736	-0.731
	(0.245)	(0.245)	(0.245)	(0.245)	(0.435)	(0.437)
Good	-0.593	-0.585	-0.591	-0.583	-0.585	-0.582
	(0.243)	(0.243)	(0.243)	(0.243)	(0.435)	(0.437)
Very Good	-0.742	-0.733	-0.740	-0.731	-0.675	-0.676
	(0.247)	(0.247)	(0.247)	(0.247)	(0.438)	(0.440)
Excellent	-0.591	-0.582	-0.588	-0.579	-0.726	-0.727
	(0.250)	(0.250)	(0.250)	(0.251)	(0.444)	(0.445)
Employment		-0.0519		-0.0521		
		(0.0625)		(0.0625)		
Hard Drug Use			0.0599	0.0602		
			(0.0864)	(0.0864)		
Log of wage						0.0724
						(0.0434)
Control	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y	Y	Y
Observations	14,047	14,047	14,047	14,047	5,085	5,085
Number of individuals	1,404	1,404	1,404	1,404	718	718

Note: Robust standard errors in parentheses. The estimated coefficients are shown in the table. Violent behavior is defined from the question "Have you attacked someone with the idea of seriously hurting them or have had a situation end up in a serious fight or assault of some kind?" Year FE for logit models are two-year dummies. Control variables are a female dummy, race dummies, age, age squared, years of education, and marital status.

with wages at the top or bottom 0.5%. Table A.6 shows the summary statistics of the sample.

Table A.5: Sample selection (PSID)

Reason of dropping sample	Dropped	Total
Household heads, aged 22 - 65		79682
Missing (Health)	320	79362
Missing (Race)	86	79276
Missing (Education)	3931	75345
Missing (Marital status)	5	75340
Missing (Employment status)	1656	73684
Summary Statistics		73684
Missing (forward employment status)	17991	55693
Regression (Employment)		55693
Missing wage	14454	59230
Wages (top and bottom 0.5%)	592	58638
Missing (forward wage)	18290	40348
Regression (Log wage)		40348

A.2.2 Employment regression

The regression model for employment is as follows:

$$\Pr\left(emp_{it+1}\right) = G\left(\alpha_0^e + \alpha_1^e HealthDummy_{it} + \alpha_X^e X_{it}^l\right), \tag{A.1}$$

where emp_{it+1} is a dummy variable which equals one if individual i is employed in survey year $t+1$ and zero otherwise,[40] and G is the logistic function. $HealthDummy_{it}$ is a vector of dummy variables for each health status, and we set the worst health status Poor as the base level. X_{it}^l is a vector of other control variables.[41] Control variables in X_{it}^l include a female dummy, race dummies for black and others, age, squared age, education category dummies (high school diploma, some college, college graduate, more than college), and marital status.

Table A.7 shows the regression result of employment status. The first to the third columns in Table A.7 show the result based on the logit model. The results in the first two columns show that all of the coefficients of health dummies take positive values, which shows that, relative to the Poor health status, better health is associated with higher employment probability. Moreover, the estimated coefficients generally increase as health status improves. The third column is the result of logit regression with a different health measure: Good Health takes one if health is either

[40] The employment status is based on the status at the interview date in survey period $t+1$. Since PSID is a biannual survey, emp_{it+1} is the employment status two years after the survey year t.

[41] We consider also the year fixed effects and the individual fixed effects in some specifications.

Table A.6: Summary statistics (PSID)

N=73684	Mean	S.D.	Min	Max	Poor	Fair	Good	Very good	Excellent
Age	43.3	12.0	22	65	50.8	46.8	44.6	42.5	40.3
Female	0.28	0.45	0	1	0.40	0.39	0.32	0.26	0.20
White	0.78	0.41	0	1	0.73	0.65	0.74	0.83	0.83
Black	0.16	0.37	0	1	0.21	0.28	0.20	0.12	0.12
Other race	0.055	0.23	0	1	0.058	0.070	0.061	0.046	0.053
Married	0.48	0.50	0	1	0.34	0.35	0.44	0.51	0.56
Health	3.61	1.04	1	5	1	2	3	4	5
Years of education	13.5	2.48	1	17	11.7	12.4	13.2	13.9	14.2
Employed	0.81	0.40	0	1	0.25	0.58	0.80	0.87	0.90
Unemployed	0.066	0.25	0	1	0.078	0.11	0.074	0.056	0.048
Out-of-labor-force	0.13	0.33	0	1	0.67	0.31	0.13	0.071	0.052
Wage	18.5	33.0	0.00031	3287.4	12.6	13.1	16.6	19.5	21.9
Log wage	2.60	0.79	-8.08	8.10	2.17	2.28	2.51	2.66	2.75
Hours of work	2072.9	733.6	1	5824	1571.0	1836.8	2022.2	2111.9	2191.1

Table A.7: Employment regression on health status (PSID)

Employment	(1) Logit	(2) Logit	(3) Logit	(4) DID	(5) SGMM
Fair	1.386	0.669			
	(0.0719)	(0.115)			
Good	2.215	0.910			
	(0.0688)	(0.117)			
Very Good	2.511	1.025			
	(0.0693)	(0.121)			
Excellent	2.646	1.115			
	(0.0726)	(0.130)			
Good Health			0.373		0.0895
			(0.0577)		(0.0976)
Negative Health Shock				-0.0513	
				(0.00924)	
Control	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Individual FE	N	Y	Y	Y	Y
Observations	55,693	22,195	22,195	40,049	43,372
Number of individuals		3,486	3,486	9,049	9,442

Note: Robust standard errors in parentheses. The second and third columns for logit results are the effects on the log of odds ratio. See Appendix B for details about SGMM. For the SGMM regression, one period lagged health is used as an instrumental variable.

Good, Very Good, or Excellent, and zero otherwise. Similar to the second column, better health is positively associated with employment status.

There could be potential endogeneity problems.[42] To mitigate the endogeneity problem, we employ a difference-in-difference (DID) approach and the system generalized method of moments (SGMM) specification.[43] The DID result in the fourth column indicates that a negative health shock decreases the employment probability by 5% points. The last column in Table 1 is based on the SGMM model. Good health status is positvely associated with the employment probability. We conclude that the regression results support the relationship between health and employment status.

[42] Existing literature on health and the labor market has focused on potential endogeneity problems. For example, better employment outcomes enable people to access better healthcare.

[43] See Appendix B for details of these empirical models.

A.2.3 Wage regression

Similarly, the regression equation for the wage rate is given by[44]

$$\ln w_{it+1} = \alpha_0^w + \alpha_1^w HealthDummy_{it} + \alpha_X^w X_{it}^l + \varepsilon_{it}^w. \tag{A.2}$$

Wage regression results are shown in Table A.8. The first and second columns are the results of OLS regression with different health measurements. The coefficients for the health status are positive and statistically significant, and the magnitude is increasing as health improves. The result shows that better health is associated with a higher wage rate, consistent with much of the literature, such as García-Gómez (2011) and García-Gómez et al. (2013).

Endogeneity problems of the wage regression have been studied extensively in the literature.[45] While Individual fixed effects control for the time-invariant characteristics of each individual, time-variant factors could be a source of the endogeneity. To further mitigate the endogeneity problem of health, we perform regressions using DID and SGMM specifications. In the DID approach, we measure the effect of a negative health shock among those who have a good health (Good, Very Good, or Excellent), comparing those who have got a negative health shock and those who have not. The result shows that, when a worker suffers a negative health shock, the wage rate decreases by 2%. In the last column, we also perform SGMM to solve the simultaneity and persistence of wage and health. The result in the fourth column supports that good health increases wages. In summary, we find that better health is positively associated with the wage rate.

[44] Note that in PSID, health status is asked at the time of the interview, while the wage rate is calculated from the salary in the last year.

[45] Although one potential cause of the endogeneity is the effect of wage on health, this reverse causality is not a problem since one-period forward wage is used as the dependent variable. Besides, the effect of wages on health is not apparent. For example, Cai (2009) uses household survey data in Australia and assumes that health and wage are simultaneously determined. She finds a significant effect of health on wages, but the opposite direction is insignificant.

Table A.8: Wage regression on health status (PSID)

Log wage	(1) Panel	(2) Panel	(3) DID	(4) SGMM
Fair	0.0394			
	(0.0301)			
Good	0.0630			
	(0.0302)			
Very Good	0.0629			
	(0.0306)			
Excellent	0.0555			
	(0.0311)			
Good Health		0.0260		0.194
		(0.0112)		(0.0901)
Negative Health Shock			-0.0235	
			(0.0142)	
Control	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y
Observations	40,348	40,348	30,324	29,084
Number of individuals	9,862	9,862	7,918	7,466

Note: Robust standard errors in parentheses.The sample does not contain observations with the top and bottom 0.5% of the wage. See Appendix B for details about SGMM.

Table B.9: **Average semi-elasticity of crime with respect to health**

Average semi-elasticity	(1) Property	(2) Property	(3) Violent	(4) Violent
Fair	-0.477	-0.635	-0.753	-0.529
	(0.219)	(0.293)	(0.180)	(0.238)
Good	-0.837	-0.821	-1.068	-0.576
	(0.211)	(0.291)	(0.173)	(0.236)
Very Good	-0.910	-0.845	-1.433	-0.721
	(0.211)	(0.294)	(0.174)	(0.240)
Excellent	-1.127	-0.717	-1.418	-0.574
	(0.212)	(0.297)	(0.174)	(0.243)
Observations	82776	14391	82776	14047

Note: Standard errors in parentheses. The observation is the number used to estimate the coefficients. To derive the marginal effects for the model with fixed effects, we use the sample mean of all observations (82776).

B Empirical models

B.1 Average semi-elasticity of crime with respect to health

In the empirical results, we show the estimated coefficients of the logit regressions. However, one would be interested in the marginal effect of the health on crime. To this end, we derive the average semi-elasticity of crime with respect to health:

$$\text{Average semi-elasticity} = E\left(\frac{\partial \ln \Pr\left(crime = 1 | health, \mathbf{X}_{it}^{c}, \eta_i, \eta_t\right)}{\partial health}\right)$$

Table B.9 shows the average semi-elasticity based on the method introduced by Kitazawa (2012).

B.2 System generalized method of moments (SGMM)

The SGMM column in the regression results shows the result based on the system generalized method of moments. In the baseline specification, we have used the forward variables of the outcomes one period ahead. However, the regression model ignores the dynamic interaction between health and crime. Now it is well-documented that the estimates of a dynamic panel model with fixed effects are inconsistent. The SGMM addresses this issue by controling for the past values of the outcome variable and taking advantage of the levels and difference equations.

Let y be an outcome variable. We consider the following model:

$$y_{it} = \gamma_1 y_{it-1} + \gamma_2 y_{it-2} + \beta GoodHealth_{it} + \alpha_{en} X_{it}^{en} + \alpha_{ex} X_{it}^{ex} + \eta_i + \eta_t + \epsilon_{it}. \quad \text{(B.1)}$$

The current value of an outcome depends on its past values (due to persistence) and health status at t as well as other covariates X_{it}^{en} and X_{it}^{ex}. This model allows endogeneity between health and outcomes.

For finite t, the fixed effect estimator is inconsistent. Hence, we use the System GMM (SGMM) estimation developed by Blundell and Bond (1998).[46]

The SGMM estimators consist of the levels and difference equations. Let $\Delta y_{it} = y_{it} - y_{it-1}$. Taking the difference of (B.1) to eliminate an individual fixed effect, we get the following difference equation:

$$\Delta y_{it} = \gamma_1 \Delta y_{it-1} + \gamma_2 \Delta y_{it-2} + \beta \Delta GoodHealth_{it} + \alpha_{en} \Delta X_{it}^{en} + \alpha_{ex} \Delta X_{it}^{ex} + \Delta\eta_t + \Delta\epsilon_{it}. \quad \text{(B.2)}$$

For the estimation, we use the moment conditions following Arellano and Bond (1991): $\forall s > 2$,

$$E(y_{it-(s+2)} \Delta\epsilon_{it}) = E(GoodHealth_{it-s} \Delta\epsilon_{it}) = E(X_{it-(s-1)}^{en} \Delta\epsilon_{it}) = E(X_{it}^{ex} \Delta\epsilon_{it}) = 0 \quad \text{(B.3)}$$

Blundell and Bond (1998) introduced additional moment conditions to improve precision of the estimator: $\forall s > 2$,

$$E(\Delta y_{it-(s+2)}(\eta_i + \epsilon_{it})) = E(\Delta GoodHealth_{it-s}(\eta_i + \epsilon_{it})) = 0 \quad \text{(B.4)}$$

$$E(\Delta X_{it-(s-1)}^{en}(\eta_i + \epsilon_{it})) = E(\Delta X_{it}^{ex}(\eta_i + \epsilon_{it})) = 0 \quad \text{(B.5)}$$

The moment conditions indicate that the differences are used as instruments for the level equations.

In summary, the moment conditions for SGMM are (B.3), (B.4), and (B.5). The instruments for the level equations are lagged differences and the instruments for the difference equations are lagged levels. For the crime regression, X_{it}^{en} contains a drug use dummy and an employment dummy.

[46] For the implementation, we use xtabond2 in Stata. For further information about xtabond2, see Roodman (2009).

C Derivations

C.1 Proof of Proposition 1

Condition 3 *Assume the following regularity conditions, which we will make clear below.*

1. $1 - s \geq \rho_y$
2. φ_{ESHI} *is sufficiently small*
3. b *is sufficiently larger than* x
4. $F^0(h'|h)$ *is sufficiently close to* $F^1(h'|h)$

(i) Note that $\forall \omega \in \{0,1\}$,

$$\begin{aligned}
\tilde{W}(h,y;\omega) - \tilde{U}(h,y;\omega) &= \int_{h_{\min}}^{h_{\max}} \left[\begin{array}{c} (1-s)\left(\begin{array}{c} \varphi_{ESHI} \max\left\{V^W(h',y;\omega) - V^{NE}(h',y;\omega),0\right\} \\ +(1-\varphi_{ESHI})\max\left\{V^W(h',y;\omega) - V^{NE}(h',y;\omega),0\right\} \end{array}\right) \\ -\rho_y \left(\begin{array}{c} \varphi_{ESHI} \max\left\{V^W(h',y;2) - V^{NE}(h',y;\omega),0\right\} \\ +(1-\varphi_{ESHI})\max\left\{V^W(h',y;\omega) - V^{NE}(h',y;\omega),0\right\} \end{array}\right) \end{array} \right] dF^\omega(h'|h) \\
&= \int_{h_{\min}}^{h_{\max}} \left[\begin{array}{c} (1-\varphi_{ESHI})\left(1-s-\rho_y\right)\max\left\{V^W(h',y;\omega) - V^{NE}(h',y;\omega),0\right\} \\ +\varphi_{ESHI} \left[\begin{array}{c} (1-s)\max\left\{V^W(h',y;\omega) - V^{NE}(h',y;\omega),0\right\} \\ -\rho_y \max\left\{V^W(h',y;2) - V^{NE}(h',y;\omega),0\right\} \end{array}\right] \end{array} \right] dF^\omega(h'|h)
\end{aligned} \tag{C.1}$$

Now if $1 - s \geq \rho_y$ and φ_{ESHI} is sufficiently small such that

$$0 \leq \varphi_{ESHI} \leq \frac{\max\left\{V^W(h,y;\omega) - V^{NE}(h,y;\omega),0\right\}\left[1-s-\rho_y\right]}{\left\{ \begin{array}{c} \left[\rho_y \max\left\{V^W(h,y;2) - V^{NE}(h,y;\omega),0\right\} - (1-s)\max\left\{V^W(h,y;\omega) - V^{NE}(h,y;\omega),0\right\}\right] \\ +\max\left\{V^W(h,y;\omega) - V^{NE}(h,y;\omega),0\right\}\left[1-s-\rho_y\right] \end{array} \right\}}$$

then $\tilde{W}(h,y;\omega) - \tilde{U}(h,y;\omega) \geq 0$. Therefore, we have, $\forall \omega \in \{0,1\}$,

$$\tilde{V}^W(h,y;\omega) - \tilde{V}^U(h,y;\omega) = \varphi\left(\tilde{W}(h,y;1) - \tilde{U}(h,y;1)\right) + (1-\varphi)\left(\tilde{W}(h,y;0) - \tilde{U}(h,y;0)\right) \geq 0,$$

and

$$\tilde{V}^W(h,y;2) - \tilde{V}^U(h,y;1) = \varphi\left(\tilde{W}(h,y;2) - \tilde{U}(h,y;1)\right) + (1-\varphi)\left(\tilde{W}(h,y;2) - \tilde{U}(h,y;0)\right) \geq 0.$$

On the other hand,

$$\tilde{U}(h,y;\omega)-\tilde{N}(h,y;\omega) = \rho_y \int_{h_{\min}}^{h_{\max}} \begin{bmatrix} \varphi_{ESHI}\max\left\{V^W(h',y;2)-V^{NE}(h',y;\omega),0\right\} \\ +(1-\varphi_{ESHI})\max\left\{V^W(h',y;\omega)-V^{NE}(h',y;\omega),0\right\} \end{bmatrix} dF^{\omega}(h'|h) \geq 0, \tag{C.2}$$

and $\forall\omega\in\{0,1\}$,

$$\tilde{V}^U(h,y;\omega)-\tilde{V}^N(h,y;\omega)=\varphi\left(\tilde{U}(h,y;1)-\tilde{N}(h,y;1)\right)+(1-\varphi)\left(\tilde{U}(h,y;0)-\tilde{N}(h,y;0)\right)\geq 0.$$

Finally, it is clear that if b is sufficiently larger than x, then we have $\forall\omega,\ V^{NE}(h,y;\omega)\geq V^P(h,y)$. Hence,

$$\begin{aligned}\tilde{V}^N(h,y;1)-\tilde{V}^P(h,y) &= \int_{h_{\min}}^{h_{\max}} \begin{bmatrix} \varphi V^{NE}(h',y;1)+(1-\varphi)V^{NE}(h',y;0) \\ -\delta\begin{pmatrix}\varphi V^{NE}(h',y;1) \\ +(1-\varphi)V^{NE}(h',y;0)\end{pmatrix}-(1-\delta)V^P(h',y)\end{bmatrix} dF^1(h'|h) \\ &= (1-\delta)\int_{h_{\min}}^{h_{\max}}\begin{bmatrix}\varphi V^{NE}(h',y;1) \\ +(1-\varphi)V^{NE}(h',y;0)-V^P(h',y)\end{bmatrix} dF^1(h'|h) \\ &\geq 0.\end{aligned} \tag{C.3}$$

Also,

$$\begin{aligned}\tilde{V}^N(h,y;0)-\tilde{V}^P(h,y) &= \int_{h_{\min}}^{h_{\max}}\left[\varphi V^{NE}(h',y;1)+(1-\varphi)V^{NE}(h',y;0)\right]dF^0(h'|h) \\ &\quad -\int_{h_{\min}}^{h_{\max}}\begin{bmatrix}\delta\left(\varphi V^{NE}(h',y;1)+(1-\varphi)V^{NE}(h',y;0)\right) \\ +(1-\delta)V^P(h',y)\end{bmatrix} d\left(F^0(h'|h)-F^1(h'|h)\right).\end{aligned}$$

Now if $F^0(h'|h)$ is sufficiently close to $F^1(h'|h)$ such that the second term on the right hand side is small, then we have $\tilde{V}^N(h,y;0)\geq\tilde{V}^P(h,y)$.

Hence, we have

$$\tilde{V}^W(h,y;\omega)-\tilde{V}^P(h,y)\geq\tilde{V}^U(h,y;\omega)-\tilde{V}^P(h,y)\geq\tilde{V}^N(h,y;\omega)-\tilde{V}^P(h,y)\geq 0,$$

which implies

$$\bar{m}^W(h,y;\omega)\geq\bar{m}^U(h,y;\omega)\geq\bar{m}^N(h,y;\omega)\geq 0.$$

(ii): By (C.1) and the assumption of the first-order stochastic dominance of $F^{\omega}(h'|h)$, we have

$$\tilde{V}^{W}(h_1, y; \omega) - \tilde{V}^{U}(h_1, y; \omega) \geq \tilde{V}^{W}(h_2, y; \omega) - \tilde{V}^{U}(h_2, y; \omega)$$

for any $h_1 \geq h_2$, any y, and any ω. Hence, we have

$$\frac{\partial \tilde{V}^{W}(h, y; \omega)}{\partial h} \geq \frac{\partial \tilde{V}^{U}(h, y; \omega)}{\partial h}.$$

Similarly, using (C.2) and (C.3) we have

$$\frac{\partial \tilde{V}^{U}(h, y; \omega)}{\partial h} \geq \frac{\partial \tilde{V}^{N}(h, y; \omega)}{\partial h} \geq \frac{\partial \tilde{V}^{P}(h, y)}{\partial h}.$$

Therefore,

$$\frac{\partial \bar{m}^{X}(h, y; \omega)}{\partial h} = \beta\pi \left[\frac{\partial \tilde{V}^{X}(h, y; \omega)}{\partial h} - \frac{\partial \tilde{V}^{P}(h, y)}{\partial h}\right] \geq 0$$

for $X \in \{U, W, N\}$.

For the second part, we use the following Lemma:

Lemma 4 *Suppose there is a function $\phi(h, \omega)$ satisfying*

$$\phi(h, \omega) \geq u(h, \omega) + \beta \int_{h_{\min}}^{h_{\max}} \sum_{\omega'} \tau_{\omega'}(h') \phi(h', \omega') dF^{\omega}(h'|h)$$

for some functions $u(h, \omega) \geq 0$ and $\tau_{\omega}(h) \geq 0$, then $\phi(h, \omega) \geq 0$.

Proof. *We can show it by substituting the inequality recursively:*

$$\begin{aligned} \phi(h, \omega) &\geq u(h, \omega) + \beta \int_{h_{\min}}^{h_{\max}} \sum_{\omega'} \tau_{\omega'}(h') \left[\begin{array}{c} u(h', \omega') \\ +\beta \int_{h_{\min}}^{h_{\max}} \sum_{\omega''} \tau_{\omega''}(h'') \phi(h'', \omega'') dF^{\omega'}(h''|h') \end{array} \right] dF^{\omega}(h'|h) \\ &= u(h, \omega) + \beta \int_{h_{\min}}^{h_{\max}} \sum_{\omega'} \tau_{\omega'}(h') u(h', \omega') dF^{\omega}(h'|h) \\ &\quad + \beta^2 \int_{h_{\min}}^{h_{\max}} \int_{h_{\min}}^{h_{\max}} \sum_{\omega'} \sum_{\omega''} \tau_{\omega'}(h') \tau_{\omega''}(h'') \phi(h'', \omega'') dF^{\omega'}(h''|h') dF^{\omega}(h'|h) \\ &\geq \ldots \end{aligned}$$

Since $u(h, \omega)$ and $\tau_{\omega}(h)$ are non-negative, we must have $\phi(h, \omega) \geq 0$. ■

By differentiating the value functions, we have

$$\begin{aligned}
\frac{\partial V^U(h,y;\omega)}{\partial y} &= \beta\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h,y;\omega)\right)\right)\right]\frac{\partial \tilde{V}^U(h,y;\omega)}{\partial y}+\lambda\beta\pi\left(1-H\left(\bar{m}^U(h,y;\omega)\right)\right)\frac{\partial \tilde{V}^P(h,y)}{\partial y} \\
\frac{\partial V^W(h,y;\omega)}{\partial y} &= \left(1-\tau^i\right)\frac{\partial w(h,y;\omega)}{\partial y}+\beta\left[1-\lambda\pi\left(1-H\left(\bar{m}^W(h,y;\omega)\right)\right)\right]\frac{\partial \tilde{V}^W(h,y;\omega)}{\partial y} \\
&\quad +\lambda\beta\pi\left(1-H\left(\bar{m}^W(h,y;\omega)\right)\right)\frac{\partial \tilde{V}^P(h,y)}{\partial y} \\
\frac{\partial V^N(h,y;\omega)}{\partial y} &= \beta\left[1-\lambda\pi\left(1-H\left(\bar{m}^N(h,y;\omega)\right)\right)\right]\frac{\partial \tilde{V}^N(h,y;\omega)}{\partial y}+\lambda\beta\pi\left(1-H\left(\bar{m}^N(h,y;\omega)\right)\right)\frac{\partial \tilde{V}^P(h,y)}{\partial y} \\
\frac{\partial V^P(h,y)}{\partial y} &= \beta\frac{\partial \tilde{V}^P(h,y)}{\partial y}
\end{aligned}$$

Hence, after simplifying, we have

$$\begin{aligned}
\frac{\partial V^W(h,y;\omega)}{\partial y}-\frac{\partial V^U(h,y;\omega)}{\partial y} &\geq \left(1-\tau^i\right)\frac{\partial w(h,y;\omega)}{\partial y} && \text{(C.4)} \\
&\quad +\beta\left(\left(1-\lambda\pi\left(1-H\left(\bar{m}^W(h,y;\omega)\right)\right)\right)\left[\frac{\partial \tilde{V}^W(h,y;\omega)}{\partial y}-\frac{\partial \tilde{V}^U(h,y)}{\partial y}\right]\right) \\
\frac{\partial V^W(h,y;\omega)}{\partial y}-\frac{\partial V^N(h,y;\omega)}{\partial y} &\geq \left(1-\tau^i\right)\frac{\partial w(h,y;\omega)}{\partial y} && \text{(C.5)} \\
&\quad +\beta\left(\left(1-\lambda\pi\left(1-H\left(\bar{m}^W(h,y;\omega)\right)\right)\right)\left[\frac{\partial \tilde{V}^W(h,y;\omega)}{\partial y}-\frac{\partial \tilde{V}^N(h,y)}{\partial y}\right]\right) \\
\frac{\partial V^U(h,y;\omega)}{\partial y}-\frac{\partial V^N(h,y;\omega)}{\partial y} &\geq \beta\left(1-\lambda\pi\left(1-H\left(\bar{m}^U(h,y;\omega)\right)\right)\left[\frac{\partial \tilde{V}^U(h,y;\omega)}{\partial y}-\frac{\partial \tilde{V}^N(h,y)}{\partial y}\right]\right) && \text{(C.6)} \\
\frac{\partial V^N(h,y;\omega)}{\partial y}-\frac{\partial V^P(h,y)}{\partial y} &= \beta\left(1-\lambda\pi\left(1-H\left(\bar{m}^N(h,y;\omega)\right)\right)\left[\frac{\partial \tilde{V}^N(h,y;\omega)}{\partial y}-\frac{\partial \tilde{V}^P(h,y)}{\partial y}\right]\right)
\end{aligned}$$

Now consider for $\omega \in \{0,1\}$,

$$\begin{aligned}
&\frac{\partial \tilde{W}(h,y;\omega)}{\partial y}-\frac{\partial \tilde{U}(h,y;\omega)}{\partial y} \\
&= (1-\varphi_{ESHI})\left[1-s-\rho_y\right]\left(\frac{\partial V^W(h,y;\omega)}{\partial y}-\frac{\partial V^U(h,y;\omega)}{\partial y}\right)+\varphi_{ESHI}\left[\begin{array}{c}(1-s)\left(\frac{\partial V^W(h,y;\omega)}{\partial y}-\frac{\partial V^U(h,y;\omega)}{\partial y}\right) \\ -\rho_y\left(\frac{\partial V^W(h,y;2)}{\partial y}-\frac{\partial V^U(h,y;\omega)}{\partial y}\right)\end{array}\right] \\
&= \left[1-s-\rho_y\right]\left(\frac{\partial V^W(h,y;\omega)}{\partial y}-\frac{\partial V^U(h,y;\omega)}{\partial y}\right)+\varphi_{ESHI}\rho_y\left[\frac{\partial V^W(h,y;\omega)}{\partial y}-\frac{\partial V^W(h,y;2)}{\partial y}\right]
\end{aligned}$$

Hence,

$$
\begin{aligned}
& \frac{\partial \tilde{V}^W(h,y;\omega)}{\partial y} - \frac{\partial \tilde{V}^U(h,y;\omega)}{\partial y} \\
= & \int_{h_{\min}}^{h_{\max}} \left[\varphi \left(\frac{\partial \tilde{W}(h,y;1)}{\partial y} - \frac{\partial \tilde{U}(h,y;1)}{\partial y} \right) + (1-\varphi) \left(\frac{\partial \tilde{W}(h,y;0)}{\partial y} - \frac{\partial \tilde{U}(h,y;0)}{\partial y} \right) \right] dF^{\omega}(h'|h) \\
= & \rho_y \varphi_{ESHI} \left[\varphi \left(\frac{\partial V^W(h,y;1)}{\partial y} - \frac{\partial V^W(h,y;2)}{\partial y} \right) + (1-\varphi) \left(\frac{\partial V^W(h,y;0)}{\partial y} - \frac{\partial V^W(h,y;2)}{\partial y} \right) \right] \\
& + \left(1-s-\rho_y\right) \int_{h_{\min}}^{h_{\max}} \left[\varphi \left(\frac{\partial V^W(h,y;1)}{\partial y} - \frac{\partial V^U(h,y;1)}{\partial y} \right) + (1-\varphi) \left(\frac{\partial V^W(h,y;0)}{\partial y} - \frac{\partial V^U(h,y;0)}{\partial y} \right) \right] dF^{\omega}(h'|h) \\
\geq & \left(1-s-\rho_y\right)\left(1-\tau^i\right) \left[\varphi \frac{\partial w(h_{\min},y;1)}{\partial y} + (1-\varphi) \frac{\partial w(h_{\min},y;0)}{\partial y} \right] \\
& - \rho_y \varphi_{ESHI} \int_{h_{\min}}^{h_{\max}} \left[\varphi \left(\frac{\partial V^W(h,y;2)}{\partial y} - \frac{\partial V^W(h,y;1)}{\partial y} \right) + (1-\varphi) \left(\frac{\partial V^W(h,y;2)}{\partial y} - \frac{\partial V^W(h,y;0)}{\partial y} \right) \right] dF^{\omega}(h'|h) \\
& + (1-s-\rho)\beta \int_{h_{\min}}^{h_{\max}} \left[\begin{array}{c} \varphi\left(1-\lambda\pi\left(1-H\left(\bar{m}^W(h,y;1)\right)\right)\right) \left(\frac{\partial \tilde{V}^W(h,y;1)}{\partial y} - \frac{\partial \tilde{V}^U(h,y;1)}{\partial y} \right) \\ + (1-\varphi)\left(1-\lambda\pi\left(1-H\left(\bar{m}^W(h,y;0)\right)\right)\right) \left(\frac{\partial \tilde{V}^W(h,y;0)}{\partial y} - \frac{\partial \tilde{V}^U(h,y;0)}{\partial y} \right) \end{array} \right] dF^{\omega}(h'|h)
\end{aligned}
$$

Now if φ_{ESHI} is sufficiently small such that the sum of the first two terms on the right hand side are non-negative, then, by using the above Lemma, we have

$$
\frac{\partial \tilde{V}^W(h,y;\omega)}{\partial y} \geq \frac{\partial \tilde{V}^U(h,y;\omega)}{\partial y}
$$

Similarly, we also have

$$
\begin{aligned}
\frac{\partial \tilde{V}^W(h,y;\omega)}{\partial y} & \geq \frac{\partial \tilde{V}^N(h,y;\omega)}{\partial y} \geq \frac{\partial \tilde{V}^P(h,y)}{\partial y} \\
\frac{\partial \tilde{V}^U(h,y;\omega)}{\partial y} & \geq \frac{\partial \tilde{V}^P(h,y)}{\partial y}
\end{aligned}
$$

Therefore, we have

$$
\frac{\partial \bar{m}^X(h,y;\omega)}{\partial y} = \beta\pi \left[\frac{\partial \tilde{V}^X(h,y;\omega)}{\partial y} - \frac{\partial \tilde{V}^P(h,y)}{\partial y} \right] \geq 0
$$

C.2 Proof of Proposition 2

By the proof of Proposition 1, (C.4) - (C.6), and $\phi(h,y;1) > \phi(h,y;0)$, we have $\forall (\omega,\omega') \in \{(0,0)\,(1,1)\,(2,0)\,(2,1)\}$

$$
\begin{aligned}
\frac{\partial V^{W}\left(h, y ; \omega\right)}{\partial y} & \geq \frac{\partial V^{NE}\left(h, y ; \omega^{\prime}\right)}{\partial y} \\
\frac{\partial V^{W}\left(h, y ; \omega\right)}{\partial h} & > \frac{\partial V^{NE}\left(h, y ; \omega^{\prime}\right)}{\partial h}
\end{aligned}
$$

where the second inequality is obtained by differentiating the value functions in the same manner as in the proof of Proposition 1. Also, for any (h, y, ω) such that $\frac{\partial V^{U}(h,y;\omega)}{\partial h} > 0$, we have

$$
\begin{aligned}
\frac{\partial V^{U}\left(h, y ; \omega\right)}{\partial y} & \geq \frac{\partial V^{N}\left(h, y ; \omega\right)}{\partial y} \\
\frac{\partial V^{U}\left(h, y ; \omega\right)}{\partial h} & > \frac{\partial V^{N}\left(h, y ; \omega\right)}{\partial h}.
\end{aligned}
$$

Now implicitly differentiating (22) and (23), we have

$$
\frac{\partial \bar{h}_{LFP}\left(y ; \omega\right)}{\partial y} = -\frac{\frac{\partial V^{U}\left(\bar{h}_{LFP}, y ; \omega\right)}{\partial y} - \frac{\partial V^{N}\left(\bar{h}_{LFP}, y ; \omega\right)}{\partial y}}{\frac{\partial V^{U}\left(\bar{h}_{LFP}, y ; \omega\right)}{\partial \bar{h}} - \frac{\partial V^{N}\left(\bar{h}_{LFP}, y ; \omega\right)}{\partial \bar{h}}} \leq 0
$$

and

$$
\frac{\partial \bar{h}_{EMP}\left(y ; \omega, \omega^{\prime}\right)}{\partial y} = -\frac{\frac{\partial V^{W}\left(\bar{h}_{EMP}, y ; \omega\right)}{\partial y} - \frac{\partial V^{NE}\left(\bar{h}_{EMP}, y ; \omega^{\prime}\right)}{\partial y}}{\frac{\partial V^{W}\left(\bar{h}_{EMP}, y ; \omega\right)}{\partial \bar{h}} - \frac{\partial V^{NE}\left(\bar{h}_{EMP}, y ; \omega^{\prime}\right)}{\partial \bar{h}}} \leq 0.
$$

D Model details

D.1 Flow equations

Given the existence of the health thresholds, the flow equation of the distribution function of the number of employed workers with health no greater than h, where $h \in \left[\bar{h}_{EMP}(y;\omega,1), h_{\max}\right]$, can be written as

$$\tilde{e}_y(h;2) = \int_{h_{\min}}^{h_{\max}} \left\{ \begin{array}{c} e_y(h';2)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;2)\right)\right)\right](1-s)\left[F^2(h|h') - F^2\left(\bar{h}_{EMP}(y;2,1)|h'\right)\right] \\ +\sum_{\omega\in\{0,1\}} u_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)\right]\rho_y\varphi_{ESHI} \left[\begin{array}{c} \varphi^U_{PHI}\left(\begin{array}{c} F^\omega(h|h') \\ -F^\omega\left(\bar{h}_{EMP}(y;2,1)|h'\right) \end{array}\right) \\ +\left(1-\varphi^U_{PHI}\right)\left(\begin{array}{c} F^\omega(h|h') \\ -F^\omega\left(\bar{h}_{EMP}(y;2,0)|h'\right) \end{array}\right) \end{array}\right] \end{array}\right\} dh'$$

$$\tilde{e}_y(h;1) = \int_{h_{\min}}^{h_{\max}} \left\{ \begin{array}{c} \sum_{\omega\in\{0,1\}} e_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;\omega)\right)\right)\right](1-s)\varphi^W_{PHI}\left[\begin{array}{c} F^\omega(h|h') \\ -F^\omega\left(\bar{h}_{EMP}(y;1,1)|h'\right) \end{array}\right] \\ +\sum_{\omega\in\{0,1\}} u_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)\right]\rho_y\varphi^U_{PHI}\left(1-\varphi_{ESHI}\right)\left(\begin{array}{c} F^\omega(h|h') \\ -F^\omega\left(\bar{h}_{EMP}(y;1,1)|h'\right) \end{array}\right) \end{array}\right\} dh'$$

$$\tilde{e}_y(h;0) = \int_{h_{\min}}^{h_{\max}} \left\{ \begin{array}{c} \sum_{\omega\in\{0,1\}} e_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;\omega)\right)\right)\right](1-s)\left(1-\varphi^W_{PHI}\right)\left[\begin{array}{c} F^\omega(h|h') \\ -F^\omega\left(\bar{h}_{EMP}(y;0,0)|h'\right) \end{array}\right] \\ +\sum_{\omega\in\{0,1\}} u_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)\right]\rho_y\left(1-\varphi^U_{PHI}\right)\left(1-\varphi_{ESHI}\right)\left(\begin{array}{c} F^\omega(h|h') \\ -F^\omega\left(\bar{h}_{EMP}(y;0,0)|h'\right) \end{array}\right) \end{array}\right\} dh' \tag{D.1}$$

where ρ_y is the job finding rate and

$$e_y(h;\omega) = \frac{\partial \tilde{e}_y(h;\omega)}{\partial h}$$

is the density function of employed workers. Therefore, two groups of workers become employed with health less than or equal to h : (i) those employed who did not get caught committing a crime, did not separate exogenously, and drawn a health status between $\bar{h}_2(y;\omega)$ and h, and (ii) those unemployed who did not get caught committing a crime, met with a vacant firm, and drawn a health status between $\bar{h}_{EMP}(y;\omega,\omega')$ and h.

Similarly, the flow equation of the distribution function of the number of unemployed workers with health no greater than h, where $h \in \left[\bar{h}_{LFP}(y;\omega), h_{\max}\right]$, is given by

$$
\begin{aligned}
\tilde{u}_y(h;1) = \int_{h_{\min}}^{h_{\max}} & \sum_{\omega\in\{0,1\}} e_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;\omega)\right)\right)\right]\varphi^W_{PHI}\left[\begin{array}{c} s\left(F^\omega(h|h')-F^\omega\left(\bar{h}_{LFP}(y;1)|h'\right)\right) \\ +(1-s)\left(\begin{array}{c} F^\omega\left(\min\left\{h,\max\left\{\bar{h}_{LFP}(y;1),\bar{h}_{EMP}(y;1,1)\right\}\right\}|h'\right) \\ -F^\omega\left(\bar{h}_{LFP}(y;1)|h'\right)\end{array}\right)\end{array}\right] \\
&+ e_y(h';2)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;2)\right)\right)\right]\left[\begin{array}{c} s\left(F^2(h|h')-F^2\left(\bar{h}_{LFP}(y;1)|h'\right)\right) \\ +(1-s)\left(F^2\left(\min\left\{h,\max\left\{\bar{h}_{LFP}(y;1),\bar{h}_{EMP}(y;2,1)\right\}\right\}|h'\right)-F^2\left(\bar{h}_{LFP}(y;1)|h'\right)\right)\end{array}\right] \\
&+ \sum_{\omega\in\{0,1\}} u_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)\right]\varphi^U_{PHI}\left[\begin{array}{c} \rho_y\varphi_{ESHI}\left(\begin{array}{c} F^\omega\left(\min\left\{h,\max\left\{\bar{h}_{LFP}(y;1),\bar{h}_{EMP}(y;2,1)\right\}\right\}|h'\right) \\ -F^\omega\left(\bar{h}_{LFP}(y;1)|h'\right)\end{array}\right) \\ +\rho_y(1-\varphi_{ESHI})\left(\begin{array}{c} F^\omega\left(\min\left\{h,\max\left\{\bar{h}_{LFP}(y;1),\bar{h}_{EMP}(y;1,1)\right\}\right\}|h'\right) \\ -F^\omega\left(\bar{h}_{LFP}(y;1)|h'\right)\end{array}\right) \\ +\left(1-\rho_y\right)\left(F^\omega(h|h')-F^\omega\left(\bar{h}_{LFP}(y;1)|h'\right)\right)\end{array}\right] \\
&+ \sum_{\omega\in\{0,1\}} n_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^N(h',y;\omega)\right)\right)\right]\varphi^N_{PHI}\left(F^\omega(h|h')-F^\omega\left(\bar{h}_{LFP}(y;1)|h'\right)\right)\} \\
&+ p_y(h')\,\delta\varphi^P_{PHI}\left(F^1(h|h')-F^1\left(\bar{h}_{LFP}(y;1)|h'\right)\right)\}dh' \\
\tilde{u}_y(h;0) = \int_{h_{\min}}^{h_{\max}} & \sum_{\omega\in\{0,1\}} e_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;\omega)\right)\right)\right]\left(1-\varphi^W_{PHI}\right)\left[\begin{array}{c} s\left(F^\omega(h|h')-F^\omega\left(\bar{h}_{LFP}(y;0)|h'\right)\right) \\ +(1-s)\left(\begin{array}{c} F^\omega\left(\min\left\{h,\max\left\{\bar{h}_{LFP}(y;0),\bar{h}_{EMP}(y;0,0)\right\}\right\}|h'\right) \\ -F^\omega\left(\bar{h}_{LFP}(y;0)|h'\right)\end{array}\right)\end{array}\right] \\
&+ \sum_{\omega\in\{0,1\}} u_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)\right]\left(1-\varphi^U_{PHI}\right)\left[\begin{array}{c} \rho_y\varphi_{ESHI}\left(\begin{array}{c} F^\omega\left(\min\left\{h,\max\left\{\bar{h}_{LFP}(y;0),\bar{h}_{EMP}(y;2,0)\right\}\right\}|h'\right) \\ -F^\omega\left(\bar{h}_{LFP}(y;0)|h'\right)\end{array}\right) \\ +\rho_y(1-\varphi_{ESHI})\left(\begin{array}{c} F^\omega\left(\min\left\{h,\max\left\{\bar{h}_{LFP}(y;0),\bar{h}_{EMP}(y;0,0)\right\}\right\}|h'\right) \\ -F^\omega\left(\bar{h}_{LFP}(y;0)|h'\right)\end{array}\right) \\ +\left(1-\rho_y\right)\left(F^\omega(h|h')-F^\omega\left(\bar{h}_{LFP}(y;0)|h'\right)\right)\end{array}\right] \\
&+ \sum_{\omega\in\{0,1\}} n_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^N(h',y;\omega)\right)\right)\right]\left(1-\varphi^N_{PHI}\right)\left(F^\omega(h|h')-F^\omega\left(\bar{h}_{LFP}(y;0)|h'\right)\right) \\
&+ p_y(h')\,\delta\left(1-\varphi^P_{PHI}\right)\left(F^1(h|h')-F^1\left(\bar{h}_{LFP}(y;0)|h'\right)\right)\}dh'
\end{aligned}
\tag{D.2}
$$

where

$$u_y(h;\omega) = \frac{\partial \tilde{u}_y(h;\omega)}{\partial h}$$

is the density function of the unemployed workers.

Moreover, the flow equation of the distribution function and density function of the number of workers not in the labor force with health no greater than h, where $h \in \left[h_{\min}, \bar{h}_{LFP}(y;1)\right]$, is given by

$$
\begin{aligned}
\tilde{n}_y(h;1) &= \int_{h_{\min}}^{h_{\max}} \left\{ \begin{array}{c}
\sum_{\omega\in\{0,1\}} e_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;\omega)\right)\right)\right]\varphi^W_{PHI}\left[\begin{array}{c} sF^{\omega}(h|h') \\ +(1-s)F^{\omega}\left(\min\left\{h,\bar{h}_{EMP}(y;1,1)\right\}|h'\right)\end{array}\right] \\
+e_y(h';2)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;2)\right)\right)\right]\left[\begin{array}{c} sF^{\omega}(h|h') \\ +(1-s)F^{\omega}\left(\min\left\{h,\bar{h}_{EMP}(y;2,1)\right\}|h'\right)\end{array}\right] \\
+\sum_{\omega\in\{0,1\}} u_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)\right]\varphi^U_{PHI}\left[\begin{array}{c} \rho_y\varphi_{ESHI}F^{\omega}\left(\min\left\{h,\bar{h}_{EMP}(y;2,1)\right\}|h'\right) \\ +\rho_y(1-\varphi_{ESHI})F^{\omega}\left(\min\left\{h,\bar{h}_{EMP}(y;1,1)\right\}|h'\right) \\ +\left(1-\rho_y\right)F^{\omega}(h|h')\end{array}\right] \\
+\sum_{\omega\in\{0,1\}} n_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^N(h',y;\omega)\right)\right)\right]\varphi^N_{PHI}F^{\omega}(h|h') \\
+p_y(h')\,\delta\varphi^P_{PHI}F^1(h|h')
\end{array}\right\} dh' \\
\tilde{n}_y(h;0) &= \int_{h_{\min}}^{h_{\max}} \left\{ \begin{array}{c}
\sum_{\omega\in\{0,1\}} e_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^W(h',y;\omega)\right)\right)\right]\left(1-\varphi^W_{PHI}\right)\left[\begin{array}{c} sF^{\omega}(h|h') \\ +(1-s)F^{\omega}\left(\min\left\{h,\bar{h}_{EMP}(y;0,0)\right\}|h'\right)\end{array}\right] \\
\sum_{\omega\in\{0,1\}} u_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)\right]\left(1-\varphi^U_{PHI}\right)\left[\begin{array}{c} \rho_y\varphi_{ESHI}F^{\omega}\left(\min\left\{h,\bar{h}_{EMP}(y;2,0)\right\}|h'\right) \\ +\rho_y(1-\varphi_{ESHI})F^{\omega}\left(\min\left\{h,\bar{h}_{EMP}(y;0,0)\right\}|h'\right) \\ +\left(1-\rho_y\right)F^{\omega}(h|h')\end{array}\right] \\
+\sum_{\omega\in\{0,1\}} n_y(h';\omega)\left[1-\pi\lambda\left(1-H\left(\bar{m}^N(h',y;\omega)\right)\right)\right]\left(1-\varphi^N_{PHI}\right)F^{\omega}(h|h') \\
+p_y(h')\,\delta\left(1-\varphi^P_{PHI}\right)F^1(h|h')
\end{array}\right\} dh' \\
n_y(h;\omega) &= \frac{\partial\tilde{n}_y(h;\omega)}{\partial h}
\end{aligned}
\tag{D.3}
$$

Finally, the flow equation of the distribution function and density function of the number of workers in prison with health no greater than h, where $h\in[h_{\min},h_{\max}]$, is given by

$$
\begin{aligned}
\tilde{p}_y(h) &= \int_{h_{\min}}^{h_{\max}} \left\{ \begin{array}{c}
\sum_{\omega\in\{0,1,2\}} e_y(h';\omega)\,\pi\lambda\left(1-H\left(\bar{m}^W(h',y;\omega)\right)\right)F^{\omega}(h|h') \\
\sum_{\omega\in\{0,1\}} u_y(h';\omega)\,\pi\lambda\left(1-H\left(\bar{m}^U(h',y;\omega)\right)\right)F^{\omega}(h|h') \\
\sum_{\omega\in\{0,1\}} n_y(h';\omega)\,\pi\lambda\left(1-H\left(\bar{m}^N(h',y;\omega)\right)\right)F^{\omega}(h|h') \\
p_y(h')(1-\delta)F^1(h|h')
\end{array}\right\} dh' \qquad \text{(D.4)} \\
p_y(h) &= \frac{\partial\tilde{p}_y(h)}{\partial h}
\end{aligned}
$$

D.2 Equilibrium and measurements

The total number of unemployed workers who are of ability y will be

$$
u_y = \sum_{\omega\in\{0,1\}} \int_{h_{\min}}^{h_{\max}} u_y(h;\omega)\,dh
$$

and hence in equilibrium the probability that a vacant firm meets with a worker with health h is given by

$$\zeta_y(h) = \frac{u_y(h;0) + u_y(h;1)}{u_y}$$

We normalize the population as one. Thus, we have

$$1 = \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} \left[\sum_{\omega \in \{0,1,2\}} e_{y_j}(h;\omega) + \sum_{\omega \in \{0,1\}} \left(u_{y_j}(h;\omega) + n_{y_j}(h;\omega) + p_{y_j}(h;\omega) \right) \right] dh$$

Also, in equilibrium, the expected criminal loss is given by

$$\begin{aligned} \tau^c \quad = \quad & \lambda \sum_{\omega \in \{0,1\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} u_{y_j}(h;\omega) \int_{\bar{m}^U(h,y_j;\omega)}^{m_{\max}} m dH(m)\, dh \\ & + \lambda \sum_{\omega \in \{0,1,2\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} e_{y_j}(h;\omega) \int_{\bar{m}^W(h,y_j;\omega)}^{m_{\max}} m dH(m)\, dh \\ & + \lambda \sum_{\omega \in \{0,1\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} n_{y_j}(h;\omega) \int_{\bar{m}^N(h,y_j;\omega)}^{m_{\max}} m dH(m)\, dh \end{aligned} \tag{D.5}$$

We can now proceed to define an equilibrium in this economy.

Definition 5 *A stationary market equilibrium is a set of cutoff crime values* $\{\bar{m}^X(h,y)\}$, *where* $X \in \{W, U, N\}$, *cutoff health statuses* $\{\bar{h}_{LFP}(y), \bar{h}_{EMP}(y)\}$, *market tightness for each market* θ_y, *distribution of individuals* $\{\tilde{e}_y(h;\omega), \tilde{u}_y(h;\omega), \tilde{n}_y(h;\omega), \tilde{p}_y(h;\omega)\}$, *and crime loss* τ^c, *such that (i) cutoff crime values satisfy (18), (ii) cutoff health statuses satisfy (22) and (23), (iii) market tightness for each market satisfies (26), (iv) distribution of individuals satisfy (D.1) - (D.4), and (v) crime loss satisfies (D.5).*

Finally, we measure the aggregate variables by summing the variables for each ability, health, and health insurance status. Therefore, the rates of unemployment, employment, non-in-the-labor-force, and in-prison are given by

$$\begin{aligned}
u &= \sum_{j=1}^{J} u_{y_j} = \sum_{\omega\in\{0,1\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} u_y\left(h;\omega\right) dh \\
e &= \sum_{\omega\in\{0,1,2\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} e_{y_j}\left(h;\omega\right) dh \\
n &= \sum_{\omega\in\{0,1\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} n_{y_j}\left(h;\omega\right) dh \\
p &= \sum_{\omega\in\{0,1\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} p_{y_j}\left(h;\omega\right) dh
\end{aligned}$$

The (observed) crime rate of the economy is given by the proportion of workers not in prison who get caught committing a crime:

$$cr = \frac{\lambda\pi \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} \left[\begin{array}{c} \sum_{\omega\in\{0,1\}} \left(1 - H\left(\bar{m}^U\left(h, y_j;\omega\right)\right)\right) u_{y_j}\left(h;\omega\right) \\ + \sum_{\omega\in\{0,1,2\}} \left(1 - H\left(\bar{m}^W\left(h, y_j;\omega\right)\right)\right) e_{y_j}\left(h;\omega\right) \\ + \sum_{\omega\in\{0,1\}} \left(1 - H\left(\bar{m}^N\left(h, y_j;\omega\right)\right)\right) n_{y_j}\left(h;\omega\right) \end{array}\right] dh}{\sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} \left(\sum_{\omega\in\{0,1\}} u_{y_j}\left(h;\omega\right) + \sum_{\omega\in\{0,1,2\}} e_{y_j}\left(h;\omega\right) + \sum_{\omega\in\{0,1\}} n_{y_j}\left(h;\omega\right)\right) dh}$$

Finally, the labor force participation rate[47] and aggregate output (GDP) in the economy are given by

$$\begin{aligned}
pr &= \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} \left(\frac{\sum_{\omega\in\{0,1\}} u_{y_j}\left(h;\omega\right) + \sum_{\omega\in\{0,1,2\}} e_{y_j}\left(h;\omega\right)}{\sum_{\omega\in\{0,1\}} \left[u_{y_j}\left(h;\omega\right) + n_{y_j}\left(h;\omega\right)\right] + \sum_{\omega\in\{0,1,2\}} e_{y_j}\left(h;\omega\right)}\right) dh \\
GDP &= \sum_{\omega\in\{0,1,2\}} \sum_{j=1}^{J} \int_{h_{\min}}^{h_{\max}} \phi\left(h, y;\omega\right) e_{y_j}\left(h;\omega\right) dh
\end{aligned}$$

[47] We exclude prisoners from the working-age population, which is consistent with the definition of BLS.

E Calibration tables

Table E.1: Transition matrix of health status

(i) Without health insurance						
		To				
		Poor	Fair	Good	Very good	Excellent
From	Poor	59.2%	26.4%	9.8%	4.6%	0.0%
	Fair	4.9%	41.2%	40.9%	10.3%	2.7%
	Good	1.0%	11.3%	51.0%	29.7%	7.0%
	Very good	0.2%	1.2%	22.9%	55.8%	19.9%
	Excellent	0.0%	2.3%	7.8%	32.8%	57.1%
(ii) With health insurance						
		To				
		Poor	Fair	Good	Very good	Excellent
From	Poor	51.8%	36.3%	11.3%	0.3%	0.3%
	Fair	6.5%	52.6%	35.6%	4.4%	0.9%
	Good	0.4%	9.0%	59.3%	27.3%	4.0%
	Very good	0.1%	1.5%	20.4%	61.8%	16.2%
	Excellent	0.1%	0.4%	8.1%	30.8%	60.6%

Note: This table shows the transition matrix of health status for those with and without health insurance respectively. The transition matrix is derived from MEPS.

Table E.2: Joint population distribution of health and education

(i) Without health insurance

		Education				
		< HS	HS	Some college	College	> College
Health	Poor	0.52%	0.17%	0.08%	0.02%	0.00%
	Fair	1.32%	0.64%	0.45%	0.21%	0.22%
	Good	2.09%	1.51%	0.86%	0.45%	0.49%
	Very good	0.86%	0.98%	1.17%	0.81%	0.61%
	Excellent	0.43%	0.61%	0.43%	0.30%	0.46%

(ii) With health insurance

		Education				
		< HS	HS	Some college	College	> College
Health	Poor	1.58%	1.05%	0.54%	0.26%	0.15%
	Fair	3.77%	1.83%	2.15%	1.12%	0.82%
	Good	4.38%	5.53%	5.43%	4.32%	4.46%
	Very good	3.60%	5.80%	6.34%	8.47%	8.18%
	Excellent	1.44%	1.87%	2.55%	4.04%	4.61%

Note: This table shows the joint distribution of health, education, and health insurance coverage. The sum of the numbers in (ii) is 84.29%, which is the health insurance coverage rate. The joint distribution is derived from PSID.

Table E.3:
Relative wage rate by health and education: model vs. data

(i) Without health insurance

		Education									
		< HS		HS		Some college		College		> College	
Health	Poor	1.00	(1.00)	1.13	(1.37)	1.21	(1.76)	1.29	(0.99)	1.36	(1.44)
	Fair	1.00	(1.04)	1.13	(1.32)	1.21	(1.43)	1.40	(1.24)	1.70	(1.38)
	Good	1.00	(1.22)	1.13	(1.35)	1.40	(1.40)	1.75	(1.59)	2.04	(2.09)
	Very good	1.00	(1.26)	1.13	(1.43)	1.61	(1.63)	1.95	(2.04)	2.25	(2.35)
	Excellent	1.00	(1.34)	1.17	(1.43)	1.75	(1.63)	2.09	(1.96)	2.39	(2.22)

(ii) With public health insurance

		Education									
		< HS		HS		Some college		College		> College	
Health	Poor	1.39	(1.41)	1.71	(1.54)	1.94	(2.72)	2.15	(3.38)	2.33	(2.34)
	Fair	1.54	(1.04)	1.93	(1.60)	2.21	(1.79)	2.60	(2.79)	3.06	(2.99)
	Good	1.65	(1.38)	2.08	(1.79)	2.59	(2.07)	3.14	(3.25)	3.62	(3.71)
	Very good	1.73	(1.34)	2.20	(2.01)	2.94	(2.13)	3.51	(2.87)	4.01	(4.16)
	Excellent	1.79	(1.68)	2.34	(2.41)	3.20	(2.50)	3.79	(3.76)	4.31	(4.43)

(ii) With employer sponsored health insurance

		Education									
		< HS		HS		Some college		College		> College	
Health	Poor	1.38	(1.40)	1.99	(1.91)	2.34	(2.51)	2.64	(2.47)	2.91	(2.93)
	Fair	1.50	(1.54)	2.19	(1.87)	2.58	(2.43)	3.08	(3.10)	3.63	(3.65)
	Good	1.57	(1.75)	2.31	(2.25)	2.92	(2.49)	3.57	(3.54)	4.14	(4.06)
	Very good	1.63	(1.87)	2.41	(2.34)	3.24	(2.79)	3.91	(3.90)	4.51	(4.46)
	Excellent	1.68	(1.75)	2.52	(2.22)	3.47	(2.75)	4.16	(4.12)	4.78	(4.91)

Note: This table shows the relative wage rate in the model and in the data. All wages are relative to the wage rate of less than high school, poor health status, and without health insurance. Wage rates from the data are shown in parentheses. All wage data are from PSID.

F Additional policies

F.1 Income tax

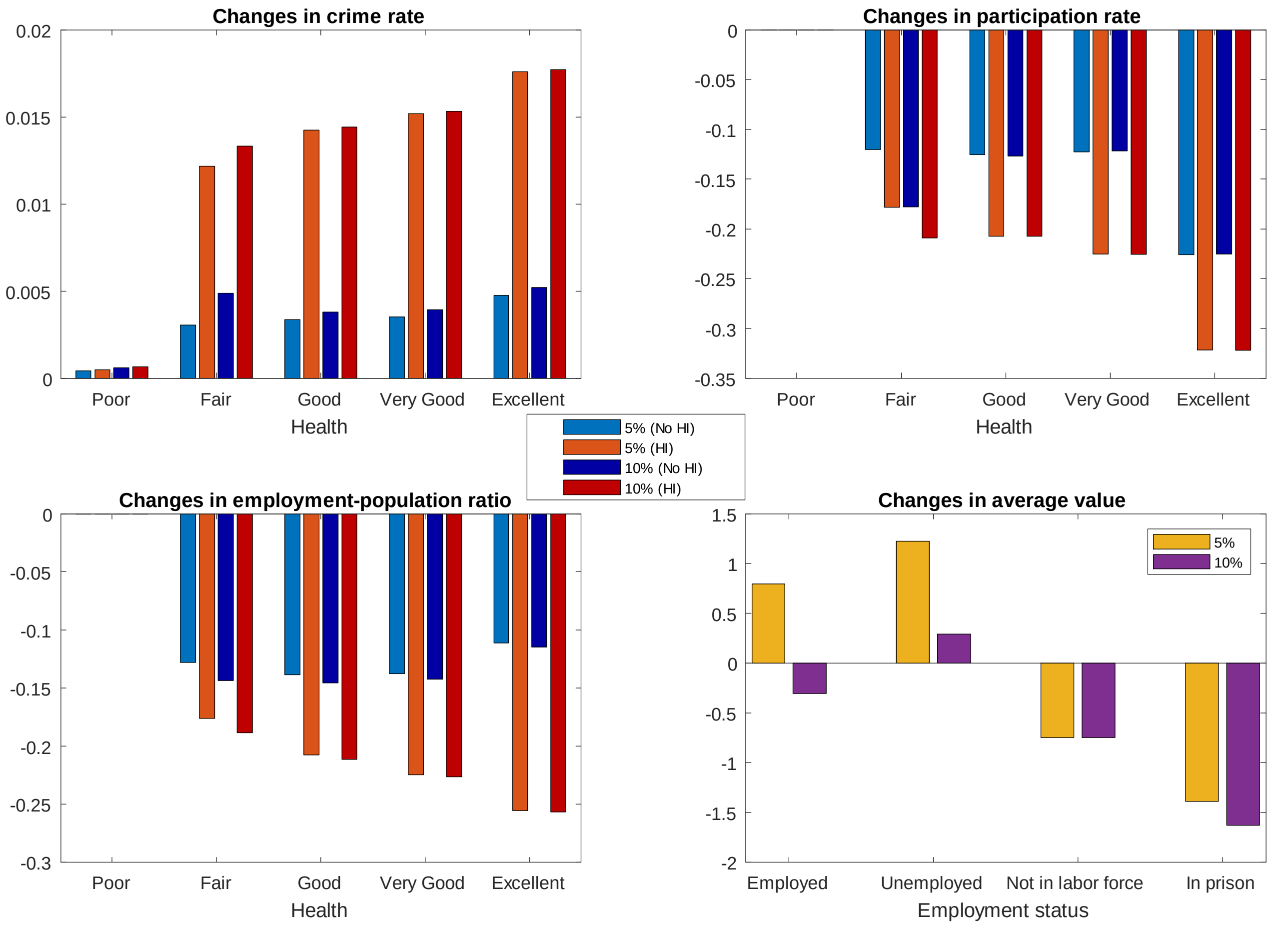

Figure F.1: Effects of income tax

Here we consider income tax policy.[48] Here an exogenous income tax τ^i is imposed to all employed workers. Hence, the value of being employed becomes

$$V^W(h, y; \omega) = \left(1 - \tau^i\right) w(h, y; \omega) - \tau^c + \lambda K^W(h, y; \omega) + (1 - \lambda)\beta \tilde{V}^W(h, y; \omega) \tag{F.6}$$

In the baseline case, there is no income tax. Here we introduce 5% and 10% income tax in the model.[49] Figure F.1 summarizes the effects by income tax at different levels.

[48] Note that we do not consider how the government uses its revenue from the income tax. If it is used, for example, to increase unemployment insurance, the actual impact of the policy may be larger or smaller. In that case, we have to consider the impact of the combination of two policies.

[49] More care should be taken when interpreting the income tax in this exercise. Since there is no income tax in

After the introduction of income tax, the crime rate in equilibrium increases regardless of health and insurance status due to the reduction of the opportunity cost of committing a crime. Income tax decreases wage income for each individual and it reduces the benefit of employment, leading to the lower opportunity cost of crime. As in the standard models without crime, taxation on income discourages participation and decreases employment-population ratio. By introducing 10% income tax, both participation rate and employment-population ratio decrease by more than 10 percentage points. The changes are larger for the insured since the insured workers are more productive and suffered more by the tax. Note that the average values for the employed and the unemployed increase due to the compositional change, since only relatively productive individuals remain in these states.

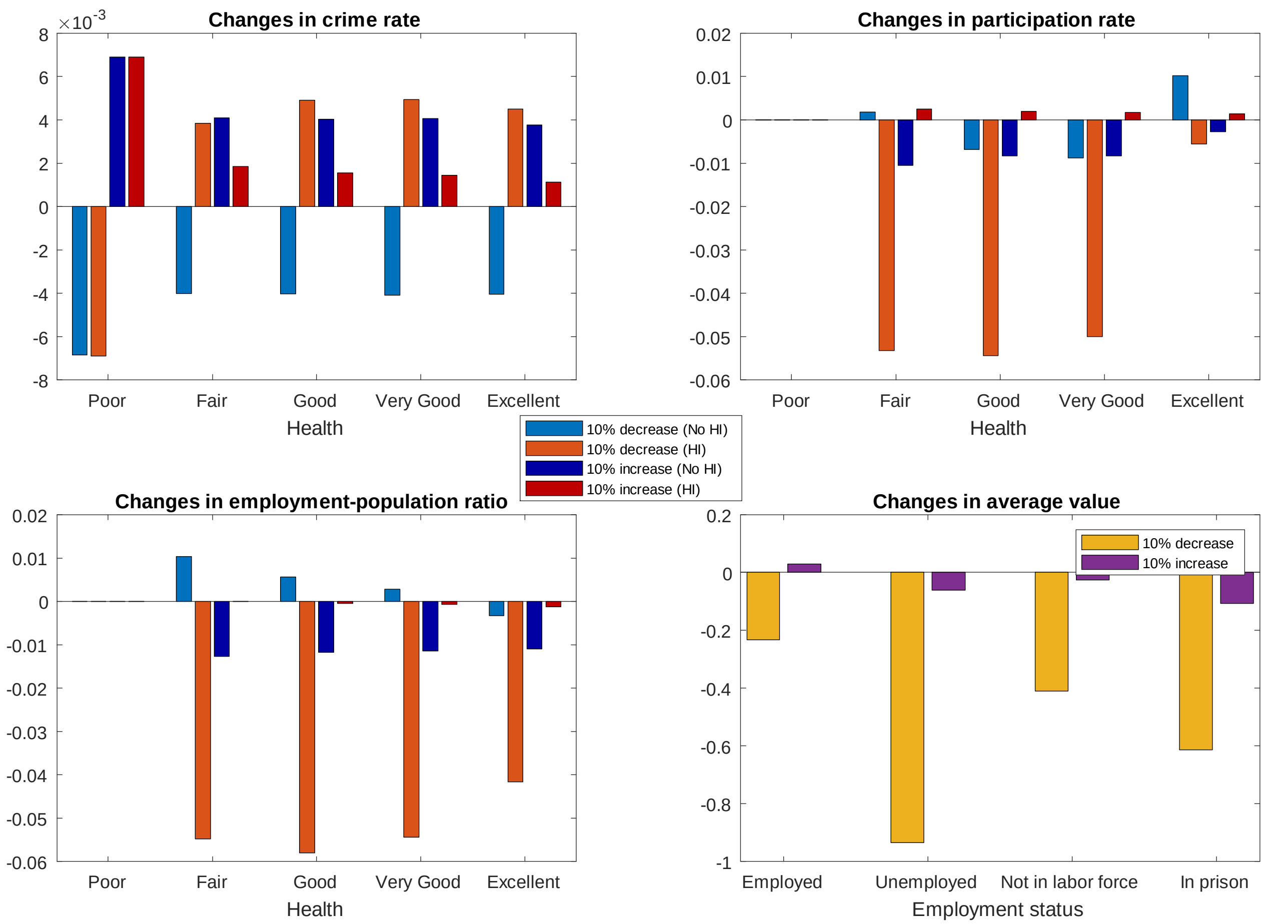

Figure F.2: Effects of changing apprehension rate

the baseline calibration, we can interpret the income tax in the model as the *additional* income tax imposed on the US economy.

F.2 Apprehension rate

Another important crime policy is to change the apprehension rate of criminals. This can be done, for example, by changing the size of the police force in a jurisdiction. Figure F.2 shows the effects of changing π, the apprehension rate of a crime. In the baseline case, we have $\pi = 0.194$. Shown in the figure are the cases when π increases by 10% ($\pi = 0.213$) and when π decreases by 10% ($\pi = 0.175$).

Intuitively, a higher apprehension rate has two effects on the (observed) crime rate. On the one hand, a higher apprehension rate would directly discourage workers from committing a crime by increasing the expected cost (the *discouragement effect*). On the other hand, a higher apprehension rate also entails that a crime, when committed by an individual, is more likely to be caught and observed (the *detection effect*). We can see that the detection effect in general dominates the discouragement effect. As a result, the crime rate increases as the apprehension rate rises. It is worth noting that the apprehension rate has only relatively small quantitative effects on labor market outcomes. In particular, the participation and employment rates are virtually unchanged under a higher value of π.