Heterogeneous Income Profiles Model with Fixed Effects: 
Incorporating Health Shocks as an Application*

Nayoung Lee†  
Chinese University of Hong Kong

Hyungsik Roger Moon‡  
USC and Yonsei University

February 14, 2015

Abstract

We provide an alternative econometrics methodology to estimate a standard heterogeneous income profiles (HIP) model. Our alternative setup allows for the HIP coefficients to be fixed in the sense that they can be arbitrarily correlated with the explanatory variables of the HIP equation and can be treated as parameters to be estimated. This fixed effects approach is more general and less restrictive than the random effects approach, which assumes that the HIP coefficients are random and exogenous. As an empirical application, we analyze how much health shocks account for the persistence and variance of income shocks by including them in a standard HIP model. Our estimation results from the Panel Study of Income Dynamics (PSID) indicate that disability shock accounts for the persistence of income shocks. The decay of the impact of income shocks to one-fifth of the initial impact takes eight years for college-educated individuals from its onset, while it takes 10 years for high school-educated individuals. Health shocks contribute up to two years to the persistence of income shocks for the college educated; income shocks from which health shocks are extracted decay to one-fifth of the initial impact in six years. Health shocks have less of an impact on the persistence of income shocks for high school-educated individuals, for whom it also takes eight and a half years for the impact of income shocks from which health shocks are extracted to drop to one-fifth of the initial impact.

JEL Classifications: C33, D31, I10, J31
Keywords: Heterogeneous Income Profiles, Fixed Effects, Income Shock, Health Shock, Disability

*We thank Jiahua Che, Fatih Guvenen, Dmytro Hryshko, Hans Van Kippersluis, Lance Lochner, Thierry Magnac, Bruce Meyer, Wallace Mok, Liugang Sheng, Hongliang Zhang, and the participants of the North American Summer Meeting, Asian Meeting of Econometric Society, and Workshop on the Estimation of Economic Models of Earnings Dynamics for their constructive comments. Special thanks to Wallace Mok for sharing his disability data. All errors are ours.

†Lee: Department of Economics, Chinese University of Hong Kong; Shatin, New Territories, Hong Kong; Email: nayoung.lee@yahoo.com.

‡Hyungsik Roger Moon: Department of Economics, University of Southern California; KAP 300, University Park Campus, Los Angeles, CA 90089-0253; Email: moonr@usc.edu.
1 Introduction

A heterogeneous income profiles (HIP) model allows all individuals to have their own income growth rates, but the literature often assumes individual-specific income growth rates to be random. This paper provides an alternative econometrics methodology to estimate a standard HIP model whose main parameters of interest are the persistence and variance of unobserved long-lasting and temporary income shocks. The alternative setup of this paper allows for heterogeneous income trends to be fixed in the sense that they can be arbitrarily correlated with the explanatory variables of the HIP equation and can be treated as parameters to be estimated.

The most widely used estimation method in the income dynamics literature is minimum distance (MD) estimation (e.g., Meghir and Pistaferri, 2004; Guvenen, 2007, 2009; Blundell, Pistaferri, and Preston, 2008; Hryshko, 2012). Assuming that the HIP coefficients are random, researchers derive a functional relationship between the empirical covariance matrix of income residuals and the parameters of the underlying model and then minimize the distance between these two to estimate the parameters. However, the existing random effects MD estimation of the HIP model has a major limitation: It does not allow for observed individual characteristics in the HIP model, which can be correlated with the random trend coefficients. If allowed, complexity arises in deriving autocovariances in the MD estimation because the correlations must be specified.

Studies often omit observed heterogeneity. Instead, they accommodate income variations due to observable time-invariant factors such as cohort, race, and education through stratification. However, the stratification strategy is restricted to having a limited number of groups of observed individual characteristics. Recently, Browning, Ejrnaes, and Alvarez (2010) extended the HIP model by allowing the HIP coefficients to be correlated with initial income in the income dynamics equation. They use the simulated method of moments with a full parametric specification of the correlated random effects instead of MD estimation.

We develop an MD estimation procedure under the fixed effects setup. The proposed fixed effect approach is more general than the random effect approach that assumes the trend components to be exogenous and uncorrelated with the other income components. Compared to the correlated random effect approach of Browning, Ejrnaes, and Alvarez (2010), ours does not specify the parametric relationship between the HIP parameters and the explanatory variables (or initial
We also illustrate the potential for empirical applications of the fixed effects HIP model that were not previously possible using a random effects approach. In particular, our framework can include observed individual characteristics as controlling variables in the HIP equation. This has a certain advantage in that we can directly control time-varying individual heterogeneity and, especially, particular exogenous events that cause income fluctuations, if observed.

The sources of income fluctuations are various and some can be observed ex post. Illness and changing jobs can be examples. Controlling for such events can make it possible to identify the contributions of particular events on income fluctuations. This is important because different types of shocks would have different degrees of persistency and also be insurable differently (Meghir and Pistaferri, 2011). In the literature, income shocks are often decomposed into permanent and transitory components, but not many studies disentangle the underlying sources of income risks. One of the exceptions is Low, Meghir, and Pistaferri (2010), who distinguish different sources of income risk and evaluate the welfare values of different insurance mechanisms. The work of Postel-Vinay and Turon (2009), Low and Pistaferri (2012), and Altonji, Vidangos, and Smith (2013) is along the lines of understanding the sources of shocks.

Fixed effects MD estimation allows us to control directly the sources of income fluctuations in an HIP model. A better understanding of the sources of income risks opens up the black box that comprises transitory and permanent income shocks and in turn allows better policy making in the income domain. In particular, we apply our fixed effect methodology to analyze a source of income risk, health shock, in the form of disability shock.

Health shocks are an often-cited driver of permanent income shocks, but to what degree has not yet been studied. We analyze how much health shocks account for the persistence and variance of income shocks by including them in the HIP model. Under certain assumptions, our estimation results indicate that disability shock somewhat accounts for the persistence of income shocks. The decay of the impact of an income shock to one-fifth of the initial impact takes eight years for college-educated individuals from its onset, without controlling for health shocks, while it takes 10 years for high school-educated individuals. Health shocks contribute up to two years to the persistence of income shocks for the college educated; income shocks from which health shocks are extracted decay to one-fifth of the initial impact in six years. Health shocks have less of an
impact on the persistence of income shocks for high school-educated individuals, for whom it also
takes eight and a half years for the impact of income shocks from which health shocks are extracted
to drop to one-fifth of the initial impact.

The remainder of the paper is organized as follows: Section 2 introduces the latent income
dynamics model; Section 3 presents the empirical methodologies for unbalanced panel data; Section
4 presents small-scale Monte Carlo simulations; Section 5 describes an application of controlling
health shocks in the HIP model using data from the Panel Study of Income Dynamics (PSID); and
Section 6 concludes the paper.

2 Model

We assume that the log income process of individual $i$ in time $t$ is generated by the following
heterogeneous income profiles model:

$$y_{it}^*(h_{it}) = f(h_{it}, \beta_i) + g(h_{it}, \gamma_t) + m(x_{it}^*, \pi) + \xi_{it}^*(h_{it}),$$

where $y_{it}^*$ is the latent log earnings, $h_{it}$ is the years of work experiences, $x_{it}^*$ is a vector of observable
characteristics. The observed log income process will be introduced later.

The function $f(h_{it}, \beta_i)$ represents the heterogeneous income profile of individual $i$ as a func-
tion of the years of experience $h_{it}$ and unobserved individual characteristics $\beta_i$. As an approximation
of $f(h_{it}, \beta_i)$, we use a $K^{th}$ order polynomial function with individual-specific coefficients:

$$f(h_{it}, \beta_i) = \beta_{i,0} + \beta_{i,1} h_{it} + \cdots + \beta_{i,K} h_{it}^K = \beta_i^t H^K(h_{it}),$$

where $\beta_i = (\beta_{i,0}, \beta_{i,1}, ..., \beta_{i,K})'$ and $H^K(h_{it}) = (1, h_{it}, ..., h_{it}^K)'$. In this paper we allow for $\beta_i$ to be
fixed in the sense that we treat $\beta_i$ as parameters to estimate. Then, $\beta_i$ can be arbitrarily correlated
with other components of income and $\beta_{i,0}$ contains any time invariant factor that affects the income.

Aggregated time effect is captured by the $g(h_{it}, \gamma_t)$, a function of the years of experience $h_{it}$
and unobserved time effect $\gamma_t$. The time effect is allowed to be age-specific. As an approximation
of \(g(h_{it}, \gamma_t)\), we also use a \(K^{th}\) order polynomial function with time-specific coefficients:

\[
g(h, \gamma_t) = \gamma_{t,0} + \gamma_{t,1} h_{it} + \cdots + \gamma_{t,L} h_{it}^K = \gamma_{t}^t H^K (h_{it}).
\]

where \(\gamma_t = (\gamma_{t,0}, \gamma_{t,1}, \ldots, \gamma_{t,K})\). The specifications of \(g\) and \(f\), thus, allow experience-earnings profiles to be a \(K^{th}\) order polynomial:\(^1\)

\[
f(h_{it}, \beta_i) + g(h_{it}, \gamma_t) = (\beta_{i,0} + \gamma_{t,0}) + (\beta_{i,1} + \gamma_{t,1}) h_{it} + \cdots + (\beta_{i,K} + \gamma_{t,K}) h_{it}^K.
\]

The function \(m(x_{it}^*, \pi)\) captures the contribution of time-varying variables, \(x_{it}^*\) on income. We assume a simple linear specification:

\[
m(x_{it}^*, \pi) = \pi' x_{it}^*.
\]

A noticeable difference from the HIP literature is that our income model can explicitly control for \(x_{it}^*\). We assume that \(x_{it}^*\) is strictly exogenous with respect to the income shock \(\xi_{it}\) but allow for any potential correlation of the individual- specific parameters, \(\beta_i\). We focus on the time-varying covariates because time-invariant characteristics can be considered parts of \(\beta_{i,0}\). The time-invariant characteristics can be also considered by stratifying on variables. For example, studies usually analyze a separate income process for each education group. However, the stratification strategy is restricted to having a limited number of groups of observed individual characteristics. It is also impossible to stratify time-varying variables. Allowing the component of \(m(x_{it}^*, \pi)\) in the HIP makes the random effects assumption difficult to maintain due to the possible correlation between \(\beta_i\) and \(x_{it}^*\). If the correlation is allowed, complexity arises in deriving autocovariances in the MD estimation because the correlations must be specified.\(^2\) Studies usually omit the component of \(m(x_{it}^*, \pi)\) and avoid such complexity. However, this cannot be a solution to the problem because the omitted \(x_{it}^*\) are included in the error.

\(^1\) Murphy and Welch (1990) argue that earnings fit better to be quartic in experience (K=4), but researchers still use the quadratic function (K=2) more often. One of the reasons is that allowing for quartic experience has negligible effects on estimated rates of return (Heckman, Lochner, and Todd, 2006). On the other hand, Baker (1997) and Guvenen (2009) use a linear approximation (K=1), arguing that the extension to be a higher order function does not noticeably affect the parameter estimates.

\(^2\) However, if there is no correlation between \(\beta_i\) and \(x_{it}^*\), then \(x_{it}^*\) can be also taken away with an additional regression.
The term $\xi_{it}^* (h_{it})$ represents unobserved income shocks of individual $i$ in time $t$ whose experience is $h_{it}$. We assume that the unobserved income shock, $\xi_{it}^* (h_{it})$, is the sum of two shocks, a long lasting shock $p_{it}^* (h_{it})$ and a temporal shock $e_{it}^* (h_{it})$,

$$\xi_{it}^* (h_{it}) = p_{it}^* (h_{it}) + e_{it}^* (h_{it}).$$

The dynamics of the long lasting component $p_{it}^* (h_{it})$ is assumed to be

$$p_{it}^* (h_{it}) = \rho p_{it-1}^* (h_{it} - 1) + \eta_{it}^* (h_{it}),$$

where $\eta_{it}^* (h_{it}) \sim iid (0, \sigma_\eta^2)$ for all $h \geq 1$ with finite higher moments. We also assume that an initial value of the persistent shock before working is zero for all individuals:

$$p_{it}^* (0) = 0.$$ \tag{1}

Notice that for $h_{it} \geq 1$, we have

$$p_{it}^* (h_{it}) = \eta_{it}^* (h_{it}) + \rho \eta_{it-1}^* (h_{it} - 1) + \cdots + \rho^{h-1} \eta_{it-h+1}^* (1).$$

The temporal shock $e_{it}^* (h_{it})$ is

$$e_{it}^* (h_{it}) = \varepsilon_{it}^* (h_{it}) + \phi \varepsilon_{it-1}^* (h_{it} - 1),$$

where $\varepsilon_{it}^* (h_{it}) \sim iid (0, \sigma_\varepsilon^2)$ for all $h_{it} \geq 1$ with finite higher moments and

$$\varepsilon_{it}^* (0) = 0 \quad \text{for all } i \text{ and } t.$$ \tag{2}

The transitory shock $e_{it}^*$ may also contain classical measurement error in income, but this study does not consider their separation.\footnote{\textsuperscript{3}The temporal shock and measurement error in many other setups are also not separately identified.} We assume that $\eta_{it}^*$ and $\varepsilon_{it}^*$ are independent so that the long lasting shocks $p_{it}^*$ and the temporal shocks $e_{it}^*$ are independent. The assumptions on the initial shocks, (1) and (2) imply that they are exogenous (i.e., non-existent or to be zero). However, we
allow for $\beta_i$ to be fixed, and it is supposed to capture individuals’ initial conditions.

If one is willing to assume that no time-varying individual heterogeneity correlated with $\beta_i$ contributes to income, allowance for possible correlations among $\beta_i$ can validate the random effects assumption on $\beta_i$. Time-varying individual heterogeneity uncorrelated with $\beta_i$ can be treated as random shocks and included in the error term $\xi_{it}^*(h_{it})$, even if omitted. The only disadvantage is that the impacts of individual heterogeneity and shocks are non-separable. The methodology proposed here, however, does not focus on the case in which one omits time-varying individual heterogeneity. We consider controlling for the time-varying variables more directly.

In practice, not many variables other than education and age are controlled for as individual demographic characteristics in the income dynamics literature. One reason is that most individual demographic characteristics are time invariant or nearly so. However, sources of income fluctuations are more various but just treated as exogenous shocks, such as job mobility, long-term unemployment, health shocks, promotions, demotions, overtime labor supply, piece-rate compensation, and bonuses and premia (Meghir and Pistaferri, 2004). Most of these are observed ex post, though the data availability depends on the survey design. It is worth using the information, if available, because controlling for such events makes it possible to identify the contributions of particular events on income fluctuations. Then, the concern is that such events are purely exogenous to time-invariant individual characteristics and individual-specific income growth.

3 Estimation Methodology

The main parameters of interest are the persistence measures of the long-lasting shock and temporary shock ($\rho$ and $\phi$, respectively), and their variances ($\sigma_n^2$ and $\sigma_e^2$, respectively). To estimate $\theta = (\rho, \phi, \sigma_n^2, \sigma_e^2)$, we use the equally weighted MD estimation method. Assuming the HIP coefficients $\beta_i$ are random and there is no observed explanatory variable $x_{it}^*$, the existing HIP MD estimation minimizes the distance between the population autocovariance function of the unknown component of the income process and its sample counterpart (Baker, 1997; Guvenen 2009).\footnote{In the literature, $m$ is often omitted, and the unknown components consist of $f$ and $\xi$ after removing $g$.}

In our set-up, since the HIP coefficients $\beta_i$ are allowed to be arbitrarily correlated with other observed characteristics in $x_{it}^*$, and the autocovariance function of the observed income process $y_{it}^*$
becomes a function of not only \( \theta \) but also unknown conditional distributions of \( \beta_i \) and \( \{x_{it}\}_{t=1,...,T} \) which are difficult to identify and estimate. To avoid such difficulty, we take an alternative approach.

The idea is to treat \( \beta_i \) as individual specific parameters to estimate. We estimate the unknown coefficients \( \beta_i, \gamma_t, \) and \( \pi \), and remove the components \( f, g, \) and \( m \) before calculating the population and sample autocovariances of \( \hat{\xi}_{it} \). To remove \( f \), we propose individual-by-individual ordinary least squares (OLS) estimations. Since the coefficients \( \beta_i, \gamma_t, \) and \( \pi \) are removed, the autocovariance function of \( \hat{\xi}_{it} \) does not depend on the unknown conditional distributions of \( \beta_i \) and \( \{x_{it}\}_{t=1,...,T} \) anymore.

A not-so-straightforward part of this approach involves deriving the population autocovariance function. This is because, when the time dimension \( T \) of the panel of income data is not sufficiently long but only the cross-sectional dimension \( N \) is large, the (fixed effect) estimator of \( \beta_i \) is not consistent and biased, while the parameters \( \gamma_t \) and \( \pi \) are consistently estimable. This implies that the sampling influence of the use of \( \hat{\beta}_i \) cannot be negligible, while asymptotically the sampling influence of the use of the estimates of \( \gamma_t \) and \( \pi \) in calculating \( \hat{\xi}_{it} \) can be negligible. We consider the sampling biases in the derivation of the population autocovariance function to equalize the biases in the sample autocovariance.

Most longitudinal income data are unbalanced, with missing observations. A further complication arises in the derivation of the population autocovariance function due to the unbalancedness of the long panel data. Our concern is that we use time-series income data available for each individual. The sampling biases can be different across individuals with different patterns of missing data and working years. We minimize the distance between the sample autocovariances and the empirical population autocovariances conditional on random missing observations to account for such a discrepancy.

### 3.1 Model for Unbalanced Panel Data during Working Years

We assume that unbalancedness of the panel is caused not only by missing observations but also by multiple cohorts with different sets of working years. We classify cohorts with the starting year of labor market experience, \( \tau \). Suppose that \( T_b \) and \( T_e \) are the beginning and final years of the entire data-available periods (across the cohorts). Let \( T(\tau) \) denote a set of working years for a cohort \( \tau \).
with an (assumed) mandatory retirement in the year $T_T$. Then, we have

$$T (\tau) = \{\max (T_b, \tau), \ldots, \min \{T_T, T_e\}\}.$$  

We set, for individual $i$ of a cohort $\tau$,

$$y_{it}^* = 0, \quad x_{it}^* = 0, \quad \xi_{it}^* = 0, \quad p_{it}^* = 0, \quad \text{and} \quad e_{it}^* = 0 \quad \text{for} \quad t \notin T (\tau).$$

For individual $i$ of a cohort $\tau$, let $s_{it}$ be the dummy that takes one if the income process $y_{it}^*$ and $x_{it}^*$ are observed for time $t \in T (\tau)$, and takes zero if not. Suppose that $y_{it}$ is the observed income process of individual $i$ of cohort $\tau$ in year $t \in T (\tau)$ with years of experience $h_{it} \equiv h_{t, \tau} = (t - \tau + 1)$ if $t \geq \tau$, and $h_{it} \equiv h_{t, \tau} = 0$ otherwise.$^5$ Then, the observed income of individual $i$ of cohort $\tau$ in year $t \in T (\tau)$ is denoted by

$$y_{it} = y_{it}^* (h_{t, \tau}) s_{it}.$$

Similarly, we define

$$x_{it} = x_{it}^* s_{it}, \quad H^K_{it} = H^K (h_{t, \tau}) s_{it}, \quad \xi_{it} = \xi_{it}^* (h_{t, \tau}) s_{it}, \quad p_{it} = p_{it}^* (h_{t, \tau}) s_{it}, \quad e_{it} = e_{it}^* (h_{t, \tau}) s_{it}.$$

Here we consider that the missing observations during the cohort specific time period $T (\tau)$ are random. That is, we assume that $s_i \equiv (s_{it} : t \in T (\tau))$ are distributed independently and identically across individuals $i$, and $s_i$ is independent of $\{p_{it}^* (h) : t \in T (\tau)\}, \{e_{it}^* (h) : t \in T (\tau)\}, \{x_{it}^* : t \in T (\tau)\}, \{\gamma^t : t \in T (\tau)\}$ and $\beta_i$ for all $i$ belonging to the cohort $I (\tau)$. This assumption is quite strong but assumed in many income dynamics papers (e.g., Guvenen, 2007 and 2009, and Hryshko, 2012).$^6$ Note also that we do not assume that $s_{it}$ are independent over $t \in T (\tau)$. We allow that $s_{it}$ and $s_{is}$ can be arbitrarily correlated for any $t \neq s$.

---

$^5$We assume $h_{it} = h_{it-1} + 1$, and $h_{it} = t - \tau + 1$ if $t \geq \tau$.

$^6$There are issues on the missing observations. The attrition propensities can be correlated with individual characteristics, and especially to the instability in earnings (Fitzgerald, Gottschalk, and Moffit, 1997). The missing can be also related with $x_{it}^*$. However, we do not consider such non-random missing in this paper because this is beyond the scope of the paper.
The observed income process of individual \( i \) of cohort \( \tau \) in time \( t \in T(\tau) \), thus, follows

\[
y_{it} = \beta^t'H_{it}K + \gamma^t'H_{it} + \pi'x_{it} + \xi_{it}
\]

\[
\eta_{it} = \eta_{it} + \epsilon_{it}.
\]

We use this model for the observed income to estimate the sample autocovariance and derive its probability limit.

### 3.2 Sample Autocovariances

We first summarize a step-by-step computation procedure to estimate the sample autocovariances for the MD objective function allowing for the HIP MD estimation with fixed effects.

**Step 1:** The time effects \( \gamma^t'H_{it}K \) are taken out. For this, we estimate \( \gamma_t \) with the time-by-time OLS regressions of \( \{y_{it}\} \) on \( \{H_{it}K\} \), then take the residuals \( \bar{y}_{it} = y_{it} - \gamma_t'H_{it}K \), where \( \gamma_t = (\sum_i H_{it}K_{it})^{-1} \sum_i H_{it}K_{it}y_{it} \). Notice that

\[
\bar{y}_{it} = y_{it} - H_{it}K\gamma_t
\]

\[
= H_{it}K\beta_i + \pi'x_{it} + \xi_{it} - H_{it}K(\hat{\gamma}_t - \gamma_t)
\]

\[
= H_{it}K\bar{\beta}_i + \bar{x}_{it}\pi + \bar{\xi}_{it},
\]

where

\[
\bar{\beta}_i = \beta_i - \left( \sum_i H_{it}K_{it}H_{it}K' \right)^{-1} \sum_i H_{it}K_{it}H_{it}K' \beta_i
\]

\[
\bar{x}_{it} = x_{it} - \left( \sum_i x_{it}H_{it}K' \right) \left( \sum_i H_{it}K_{it}H_{it}K' \right)^{-1} H_{it}K
\]

\[
\bar{\xi}_{it} = \xi_{it} - H_{it}K' \left( \sum_i H_{it}K_{it}H_{it}K' \right)^{-1} \sum_i H_{it}K\xi_{it}.
\]

Here, \( \bar{\beta}_i \) should be also dependent on time index \( t \), but we skip it since its dependence is not important for the rest of analysis.

**Step 2:** The components \( H_{it}K\bar{\beta}_i + \bar{x}_{it}\pi \) are taken out. For this, we estimate \( \bar{\beta}_i \) and \( \pi \), and then
take the residuals
\[ \hat{\xi}_{it} = \hat{y}_{it} - H_{it}^K \beta_i - \pi' \hat{x}_{it}. \]

To estimate \( \beta_i \) and \( \pi \), we use the following fixed effects estimation method. We compute \( \hat{\pi} \) first by a partitioned regression. For this, we run the individual-by-individual OLS regressions of \( \{ \hat{y}_{it} \}_{t \in T(\tau)} \) on \( \{ H_{it}^K \}_{t \in T(\tau)} \) and \( \{ \hat{x}_{it} \}_{t \in T(\tau)} \) on \( \{ H_{it}^K \}_{t \in T(\tau)} \) and take the residuals
\[
\hat{\gamma}^H \hat{y}_{it} = \hat{y}_{it} - H_{it}^K \left( \sum_{t \in T(\tau)} H_{it}^K H_{it}^{K'} \right)^{-1} \sum_{t \in T(\tau)} H_{it}^K \hat{y}_{it},
\]
and
\[
\hat{\gamma}^H \hat{x}_{it} = \hat{x}_{it} - H_{it}^K \left( \sum_{t \in T(\tau)} H_{it}^K H_{it}^{K'} \right)^{-1} \sum_{t \in T(\tau)} H_{it}^K \hat{x}_{it}.
\]

We estimate \( \pi \) by running the pooled OLS of \( \hat{\gamma}^H \hat{y}_{it} \) on \( \hat{x}_{it}^T \),
\[
\hat{\pi} = \left( \sum_{i=1}^N \sum_{t \in T(\tau)} \hat{\gamma}^H \hat{x}_{it} \hat{x}_{it}^T \right)^{-1} \sum_{i=1}^N \sum_{t \in T(\tau)} \hat{\gamma}^H \hat{x}_{it} \hat{y}_{it}.
\]

One thing to note is that \( \hat{\pi} \) is consistent when \( x_{it} \) is strictly exogenous.

Then, we run another individual-by-individual OLS regressions to compute \( \hat{\beta}_i \). Let \( \hat{\beta}_i(\pi) \) by the individual-by-individual OLS estimator of \( \{ \hat{y}_{it} - \pi' \hat{x}_{it} \}_{t \in T(\tau)} \) on \( \{ H_{it}^K \}_{t \in T(\tau)} \):
\[
\hat{\beta}_i(\pi) = \left( \sum_{t \in T(\tau)} H_{it}^K H_{it}^{K'} \right)^{-1} \sum_{t \in T(\tau)} H_{it}^K \left( \hat{y}_{it} - \hat{x}_{it}^T \hat{\pi} \right).
\]

**Step 3:** We calculate the cohort-by-cohort sample autocovariances between time \( t \) and time \( t - k \).

For each cohort \( \tau \), the sample autocovariances of the residuals for the income shocks of time \( t \) and time \( t - k \) is
\[
\hat{\gamma}(t, k; \tau) = \frac{1}{N(\tau)} \sum_{i \in \tau} \hat{\xi}_{it} \hat{\xi}_{it-k},
\]
where \( N(\tau) \) be the sizes of the cohorts \( \tau \).

---

8When \( H_{it}^{K'} = 1 \), \( \hat{\xi}_{it} \) becomes the residual of the within estimator of the conventional panel regression with the conventional fixed effects.
3.3 Population Autocovariances

The probability limits of the sample autocovariances, \( \hat{\gamma}(t, k; \tau) \), become our population autocovariances, denoted by \( \gamma(\theta, t, k; \tau) \). That is, for each cohort \( \tau \),

\[
\hat{\gamma}(t, k; \tau) \rightarrow_p \gamma(\theta, t, k; \tau)
\]

and

\[
\gamma(\theta, t, k; \tau) = \mathbb{E}[\hat{\xi}_{it}\hat{\xi}_{it-k}] = \mathbb{E}[\xi^*_{it} \xi^*_{it-k}]
\]

where \( \xi^*_{it} \) is defined below. The probability limits are derived cohort-by-cohort, and so the expectations are conditional on the cohort. We omit the notation for the simplicity.

First, note that we have \( \hat{\xi}_{it} \) instead of \( \xi_{it} \) because we consider the potential biases generated by the individual-by-individual OLS estimation to allow for the fixed effects on the HIP parameters in \( f \). Since \( \hat{\pi} \) and \( \hat{\gamma}_t \) are consistent, we can approximate \( \hat{\xi}_{it} \) as

\[
\hat{\xi}_{it} = \xi_{it} - H_{it}^{K_j} \left( \sum_{t \in T(\tau)} H_{it}^K H_i^{K_j} \right)^{-1} \sum_{s \in T(\tau)} H_{is}^K \xi_{is} + o_p(1). \tag{3}
\]

When \( T \) is short, \( \hat{\beta}_t \) is not consistent, and so the term \( H_{it}^{K_j} \left( \sum_{t \in T(\tau)} H_{it}^K H_i^{K_j} \right)^{-1} \sum_{s \in T(\tau)} H_{is}^K \xi_{is} \) is not small, and should be considered when we calculate the population autocovariances of \( \hat{\xi}_{it} \).

Also notice that

\[
\hat{\xi}_{it} = \left[ \xi_{it}^* s_{it} - H^K (h_{s, \tau})' s_{it} \left( \sum_{t \in T(\tau)} H^K (h_{t, \tau}) H^K (h_{t, \tau})' s_{it} \right)^{-1} \times \left( \sum_{s \in T(\tau)} H^K (h_{s, \tau}) \xi_{is}^* s_{is} \right) \right] + o_p(1)
\]
\[= \hat{\xi}_{it}^* s_{it} + o_p(1), \text{ say.}\]

Note that \( \xi^*_{it} \) is defined in the above equation.
Therefore, for individuals of cohort $\tau$,

$$
\hat{\xi}_{it} = s_{it} p_{it}^* - s_{it} H'_{t,\tau} \left( \sum_{t \in T(\tau)} H_{t,\tau} H'_{t,\tau} s_{it} \right)^{-1} \sum_{s \in T(\tau)} H_{s,\tau} p_{is}^* s_{is}
$$

$$
+ s_{it} e_{it}^* - s_{it} H'_{t,\tau} \left( \sum_{t \in T(\tau)} H_{t,\tau} H'_{t,\tau} s_{it} \right)^{-1} \sum_{s \in T(\tau)} H_{s,\tau} e_{is}^* s_{is}
$$

$$
= \tilde{p}_{it} + \tilde{e}_{it}, \text{ say.}
$$

Then,

$$
\gamma (\theta, t, k; \tau) = \mathbb{E} (\tilde{p}_{it} \tilde{p}_{it-k}) + \mathbb{E} (\tilde{e}_{it} \tilde{e}_{it-k}). \tag{4}
$$

Note that if the bias term $H_{it}^{Kj} \left( \sum_{t \in T(\tau)} H_{it}^{Kj} H_{jt}^{Kj} \right)^{-1} \sum_{t \in T(\tau)} H_{it}^{Kj} \xi_{it}$ could be neglected, the population autocovariances would be

$$
\gamma (\theta, t, k; \tau) = \mathbb{E} (p_{it} p_{it-k}) + \mathbb{E} (e_{it} e_{it-k}). \tag{5}
$$

We additionally define $P_{\tau}^{s_1, \ldots, s_l}$ to be the probability that the samples of time period $s_1, ..., s_l$ are observed when the individual belongs to cohort $\tau$,

$$
P_{\tau}^{s_1, \ldots, s_l} = \mathbb{P} (s_{is_1} = \ldots = s_{is_l} = 1).
$$

and

$$
f_{i,T} (t, s; \tau) = s_{it} H'_{t,\tau} \left( \sum_{t \in T(\tau)} H_{t,\tau} H'_{t,\tau} s_{it} \right)^{-1} H_{s,\tau} s_{is}.
$$
Then,

\[
\mathbb{E}(\hat{p}_t \hat{p}_{i_{t-k}}) = \sigma_n^2 \gamma_p (\rho, t, t - k; \tau) P_{t,t-k}^{\tau} \\
- \sigma_n^2 \sum_{s \in T(\tau)} \gamma_p (\rho, t, s; \tau) \mathbb{E}(f_{i,T}(t - k, s, \tau) \mid s_{it} = s_{i_{t-k}} = s_{is} = 1) P_{t,t-k,s}^{\tau} \\
- \sigma_n^2 \sum_{s \in T(\tau)} \gamma_p (\rho, t - k, s; \tau) \mathbb{E}(f_{i,T}(t, s, \tau) \mid s_{it} = s_{i_{t-k}} = s_{is} = 1) P_{t,t-k,s}^{\tau} \\
+ \sigma_n^2 \sum_{s \in T(\tau)} \sum_{w \in T(\tau)} \left[ \gamma_p (\rho, s, w; \tau) \mathbb{E}(f_{i,T}(t, s, \tau) f_{i,T}(t - k, w; \tau) \mid s_{it} = s_{i_{t-k}} = s_{is} = s_{iw} = 1) \right] P_{t,t-k,s,w}^{\tau}.
\]

and

\[
\mathbb{E}(\hat{e}_t \hat{e}_{i_{t-k}} \mid h_{it} = 1) = \sigma_e^2 \gamma_e (\phi, t, t - k; \tau) P_{t,t-k}^{\tau} \\
- \sigma_e^2 \sum_{s \in T(\tau)} \gamma_e (\phi, t, s; \tau) \mathbb{E}(f_{i,T}(t - k, s, \tau) \mid s_{it} = s_{i_{t-k}} = s_{is} = 1) P_{t,t-k,s}^{\tau} \\
- \sigma_e^2 \sum_{s \in T(\tau)} \gamma_e (\phi, t - k, s; \tau) \mathbb{E}(f_{i,T}(t, s, \tau) \mid s_{it} = s_{i_{t-k}} = s_{is} = 1) P_{t,t-k,s}^{\tau} \\
+ \sigma_e^2 \sum_{s \in T(\tau)} \sum_{w \in T(\tau)} \left[ \gamma_e (\phi, s, w; \tau) \mathbb{E}(f_{i,T}(t, s, \tau) f_{i,T}(t - k, w; \tau) \mid s_{it} = s_{i_{t-k}} = s_{is} = s_{iw} = 1) \right] P_{t,t-k,s,w}^{\tau}.
\]

where

\[
\gamma_p (\rho, t, s; \tau) = \rho^{-|t-s|} \frac{(1 - \rho^{2h_{\min(s,t),\tau}})}{1 - \rho^2} I \{ h_{\min(s,t),\tau} \geq 1 \}
\]

and

\[
\gamma_e (\phi, t, s; \tau) = \begin{cases} 
1 & \text{if } t = s = \tau \\
1 + \phi^2 & \text{if } t = s, t > \tau \\
\phi & \text{if } |t - s| = 1, \min \{t, s\} \geq \tau.
\end{cases}
\]

The detailed derivations are provided in the Appendix.

However, since the limit $\gamma (\theta, t, k, \tau)$ depends on the distribution of the missing data $s_i$, we cannot use $\gamma (\theta, t, k; \tau)$ for the population moment in the MD objective function. In particular, the probability limit $\gamma (\theta, t, k, \tau)$ contains unknown components such as $P_{\tau}^{s_1, \ldots, s_l}$ and $\mathbb{E}(f_{i,T}(t, s, \tau) \mid s_{it} = \cdots = s_{iw} = 1)$. 

14
We, thus, use an estimate of \( \gamma(\theta, t, k; \tau) \), \( \hat{\gamma}(\theta, t, k; \tau) \) in which the unknown components are replaced with their estimators. We suggest their consistent estimators in Appendix.

### 3.4 Minimum Distance Estimation

We minimize the distance between the empirical population autocovariances and its sample counterpart on a cohort-by-cohort basis. However, to reduce the number of equations, we add up these with the weights of the number of observations for each cohort so that we have \( T \times T \) equations in the MD estimation. That is, for each \((t, k)\) combination, we add up \( \hat{\gamma}(t, k; \tau) \) and \( \gamma(\theta, t, k; \tau) \) over \( \tau \) with the weights of the number of observations in \( t \) and \( t - k \) for each cohort \( \tau \). Then, we stack \( \hat{\gamma}(t, k) \) and \( \gamma(\theta, t, k) \) over \( t \) and \( k \); and we denote them as \( \hat{\gamma} \) and \( \gamma(\theta) \). The MD estimator is:

\[
\hat{\theta} = \arg \min_{\theta \in \Theta} \left( \hat{\gamma} - \gamma(\theta) \right)' W (\hat{\gamma} - \gamma(\theta)).
\]

In this paper, we use \( \hat{W} = I \), which called the equally weighted MD (EWMD) estimator. The reason to choose \( \hat{W} = I \) is to avoid the many moment problem (e.g., Altonji and Segal, 1996). We compute standard errors by the bootstrap.

### 4 Monte Carlo Simulations

In this section, we investigate the finite sample performance of our estimation method using a Monte Carlo study. The data generating process for the latent income process is\(^9\)

\[
y_{it}^* = \beta_{i,0} + \beta_{i,1} h_{it} + \beta_{i,2} h_{it}^2 + \pi x_{it}^* + p_{it}^* + e_{it}^*
\]

\[
p_{it}^* = \rho p_{it-1}^* + \eta_{it}^*
\]

where \( \rho \in \{0.75, 0.85\} \), \( \pi = 0.5 \), \( \{x_{it}\} \sim \text{i.i.d.} N(0, 0.1) \), \( \{e_{it}\} \) and \( \{\eta_{it}\} \sim \text{i.i.d.} N(0, 0.05) \), and \( p_{i0} = 0 \). In addition, we generate the HIP coefficient as

\[
\beta_{i,k} = aX_i' + u_i
\]

\(^9\)The component \( g \) is omitted from the model for the simplicity. We also assume \( \mu = 0 \), because we observe that the objective function is quite flat with the MA(1) component. Many other studies might have the same issue based on their relatively huge variance of their MA(1) terms.
where $X_i = (x_{i1}, \ldots, x_{iT})$ is a collection of $x_{it}$ for each $i$, and $a = (a_1, \ldots, a_T)$ is a vector of correlation coefficients to allow for the correlation between $\beta_{i,k}$ and $x_{it}$ and among $\beta_{i,k}$: $\{a_t\}$ and $\{u_i\} \sim i.i.d. N(0, 0.1)$. The data generating process in (6) is just to allow $\beta_i$ to be correlated with the explanatory variables, and we do not estimate $a$ nor $\sigma_u$.

We consider two MD objective functions: one with the sampling bias correction and the other without considering it. The former uses the estimates of the population covariances of (4) as they are, while the latter uses the estimates of the population covariances of (5) that the bias term $H_{it}^K \left( \sum_{t \in T(\tau)} H_{it}^K H_{it}^{K'} \right)^{-1} \sum_{s \in T(\tau)} H_{is}^K \xi_{is}$ is neglected from (3). We are particularly motivated by allowing for the HIP coefficients to be fixed and by using the individual-by-individual OLS regressions to estimate and remove the fixed effects. The concern here is the potential sampling biases in the fixed effects generated by the OLS regressions with short $T$. Then, the sampling biases can also lead to biases in the MD estimates of our main parameters of interests, $\theta = (\rho, \sigma_{\eta}^2, \sigma_{\zeta}^2)$. We correct for the biases by equalizing the autocovariances for the sample and the population, but the magnitude of the biases has not yet been analyzed. Although we do not try to theoretically analyze the impact of such sampling biases in our MD estimates, it is worth seeing the simulation results.\textsuperscript{10}

Our approach involves a further complication in the derivation of the population autocovariance function due to unbalanced panel data and so it is more important to allow for unbalancedness in this simulation study. Therefore, we replicate the PSID sample in terms of the number of person–year observations and patterns of missing data. For this, we take each individual’s experience $h_{it}$, missing data patterns $s_{it}$, and working years $T(\tau)$ from the PSID data that we use for the empirical application and generate unbalanced income profiles based on these. We fix the sample size $N = 5,070$, based on the overlapped sample construction (described later), and $T = 29$. The PSID data and sample selection are described in Section 5.

The results for the fixed effects MD estimations with the sampling bias correction are reported in Table 1. Here we consider two different data generating processes: with and without $x_{it}^*(i.e., \pi \neq 0$ vs $\pi = 0$). However, for both cases, $\beta_i = (\beta_{i,0}, \beta_{i,1}, \beta_{i,2})$ are allowed to be correlated with each other. Panels (1) and (2) show the results without $x_{it}^*$ and with $x_{it}^*$, respectively. For\textsuperscript{10} There are studies that analyze the bias in AR(1) panel estimation with fixed effects and incidental trends (e.g., Phillips and Sul, 2007).
each simulation, we report the mean of the biases, standard deviations of the estimates (which is supposed to be identical to the standard deviation for the biases), and root mean squared error (RMSE).

We find quite similar precisions for all cases, regardless of \( \rho \in \{0.75, 0.85\} \) and the cases with and without \( x_{it}^* \). The biases and their standard deviations are small.

Table 2 shows the simulation results when we ignore the sampling bias correction. Panel (1) reports the simulation results for our PSID replicated samples for \( T = 29 \). We omit \( x_{it}^* \) to show how much severe the fixed effects bias is in a simpler setting, though the result with \( x_{it}^* \) is not supposed to be different since \( \pi \) can be consistently estimated. We can observe the estimates are severely biased without the bias correction. We minimize the objective function with the constraint for the estimated variances to be positive because the variances of transitory shocks are estimated to be negative without the constraint.

Then, we also examine whether the biases can be reduced in the case of one cohort without attrition, as well as in cases with a longer \( T \). We consider \( T = 40 \) for the special case in which we observe the entire income process for the working years from 25 to 65 and even longer \( T = 100 \) for the case of monthly (or quarterly) income dynamics. For these simulations, we simplify our setting by assuming the individual component \( f \) to be a linear function of experience.

We consider the case without attrition because some observations may suffer greater attrition and their small individual-by-individual OLS regressions can cause more severe biases.

The results are reported in columns (2) and (3) of Table 2. The estimates are still not close enough to the true values, but this is consistent with the findings in the econometrics literature in terms of the negative bias in \( \rho \). When \( T = 100 \), the biases are relatively reduced but still quite different from the true values. This demonstrates that the bias correction should be carried out if one takes out the individual-specific income components through individual-by-individual OLS regressions.

---

11 All the simulation results are based on 100 repetitions. The number of repetitions is small, but we observe robust results when we increase the number of repetitions.

12 The biases in a quadratic case are found to be more serious, but we omit the results.

13 Philips and Sul (2007) analyze the bias in the AR(1) dynamic panel estimation with fixed effects and incidental trends (i.e. heterogeneous growth components in our case). The bias in the AR(1) estimate is \( \hat{\rho} - \rho \rightarrow - \frac{2(1+\rho)}{T} (1 + O(\frac{1}{T})) \).

14 One may notice that the RMSE and the mean of the bias are quite similar up to the three decimal points, but they are not the same. This is mainly because of the small variations in the biases over the simulations.
5 Empirical Application

Our setup allows for sources of income shocks observed ex post in an HIP model, even when they are correlated with unobserved individual heterogeneity in deterministic income trends. In this section, we apply our methodology to measure the persistence and variance of income shocks, controlling for individuals’ health status as an observed explanatory variable in the HIP model.

Illness can reduce one’s economic opportunities by lowering the labor supply and productivity. The incidence of illness is not entirely exogenous but highly unpredictable and is thus one of the major sources of income shocks. Incorporating health shocks into the model of income dynamics is useful in case researchers want to extract health shocks from total income shocks. This allows us to know to what extent health shocks account for the persistence and variance of income shocks.\footnote{Based on this estimation, one can also investigate the effect of health shocks on inequality or on consumption via counterfactual analysis using a life-cycle model. Our estimation can also be used to measure income insurability for insurance such as unemployment compensation or disability insurance. This is because, if there is perfect insurance, then income shock persistence should be the same regardless of whether health shock is controlled for or not.}

Importantly, health shocks are not entirely exogenous (e.g., Grossman, 1972; Currie and Madrian, 1999). In our framework, we control for the endogeneity of health shocks with respect to unobserved factors affecting income by allowing unobserved individual time-invariant heterogeneity and individual-specific income trend components to be correlated with health shocks.\footnote{Our analysis is based on the assumption that health shocks are exogenous after we allow for unknown individual-specific income components to be correlated with health shocks. This may still be a strong assumption because various sources of income risk can lead health status to becoming endogenous to income. Job displacement can be an example. Non-voluntary unemployment can also cause negative health shocks (Black, Devereux, and Salvanes, 2012). Our methodology, after all, can also control for other sources of income risk together, such as job displacement status, if observed.} This treatment on the endogeneity of health shocks relaxes the usual approach in the literature, which often assumes that health status is exogenous.\footnote{Low and Pistaferri (2012) estimate the variances of income shocks controlling for disability by assuming that health status is exogenous. Their approach is also based on the restricted income profiles model that does not allow for changes in the level of income shock persistence due to health status. They only allow for changes in the variance of income shock.}

5.1 Data

We use the PSID for the 29 years from 1968 to 1996. The sample selection mainly follows Guvenen (2009): male heads of household between ages 20 and 64 who have reported positive labor earnings and hours for 20 years (not necessarily consecutive), with certain ranges of working hours and wage rates. Our income variable $y_{it}$ as the labor earnings and education variable for group analysis is
also defined following Guvenen (2009).\textsuperscript{18}

The income profiles depend on the number of years of experience and our MD estimation is on a cohort basis. We obtain the starting year of each individual’s labor market experience ($\tau$) and calculate the individual’s potential experience from the starting year of experience as $h_{it} = t - \tau + 1$.\textsuperscript{19} One concern about the cohort basis approach is that a particular cohort can have a very small number of observations and can lead to biases. To avoid this, we group individuals whose first labor market experience is from $\tau - 2$ to $\tau + 2$, and define their cohort as $\tau$.\textsuperscript{20} We thus obtain five-year overlapped observations and our final sample size is $N = 5,070$ with 27 cohorts.\textsuperscript{21}

Regarding health shocks, we obtain annually reported binary information from the PSID, which asks whether the interviewees have any physical or nervous condition that limits the type or amount of work they can do. A disability variable is defined based on their answers:\textsuperscript{22}

\[
x_{it} = \begin{cases} 
1 & \text{if yes} \\
0 & \text{if no}
\end{cases}
\]

This disability variable is exactly the same as that of Meyer and Mok (2013).\textsuperscript{23} We control for this self-reported disability variable as a proxy for one’s health status.\textsuperscript{24}

The validity of the disability measure has been argued in the literature, even in regard to measuring the disability status itself (instead of measuring health status). Many researchers have argued its subjectivity because self-reported disability may not be the same as objective measures.

\textsuperscript{18}Labor earnings include wage income, bonuses, commissions, and the labor portions of several types of income, such as farm income and business income. The PSID collects labor earnings data retrospectively, for the previous year, and so the years covered in this exercise are from 1967 to 1995. See Guvenen for a detailed description for data.

\textsuperscript{19}This is only a proxy measure for actual experience, but construction of a measure of actual experience is also controversial (Guvenen, 2009).

\textsuperscript{20}This grouping also allows us to compare our estimates with Guvenen’s, since he also analyzes cohorts with five-year overlaps in experience for the same reason.

\textsuperscript{21}Because the earliest and latest cohorts (four each) are not exactly quintupled, $N \neq 5,890$.

\textsuperscript{22}The PSID also asks those who answer yes whether their condition keeps them from doing certain types of work and, for those who answer yes or no, how much their condition limits their work, with the options "a lot," "somewhat," "just a little," and "not at all." It is possible to use such information to construct more stratified health condition variables, but we do not consider using such information in this exercise.

\textsuperscript{23}See Meyer and Mok (2013) for a detailed data cleaning process for the disability variable.

\textsuperscript{24}Currie and Madrian (1999) describe the difficulty of measuring health status: While everyone has some idea of what is meant by the term, it is remarkably difficult to measure. We assume that a health shock can be observed ex post, in the form of a disability, without measurement error. Measurement error of the health shock data could cause difficulties in interpreting the estimation result. For example, classical measurement error, if any, in the health variable would be captured in the transitory shock and then the variance of transitory shock could be overestimated. If measurement error is persistent, it can also act as a long-lasting shock.
of disability. Individuals can also report health limitations to justify non-participation in the labor force (Bound, 1991). Admittedly, our measure of health status is disputable, but it is the best available measure and widely used in the literature (see Low and Pistaferri (2012) and Meyer and Mok (2013) for a review of the controversy; they also argue the merits of the disability variable).

Table 3 shows the disability rates for 1,178 individuals (362 college educated and 816 high school educated) who satisfy the sample selection criteria. The disability rate for all observations pooled over years is 7.6%, which is lower than the 11–15% of Meyer and Mok (2013). This is because our sample requires at least 20 years of work (out of the 29-year period). Our sample is quite selected with the restriction, but note that no additional observations are omitted due to the disability variable from our original sample selection. Individuals report more disability as they get older. High school educated individuals overall report more disability than the college educated.

Table 3 also reports the prevalence of disability. Almost 40% of all individuals report a disability at least once over the years, though the prevalence rate is less for college-educated individuals (34%). More than one-third of all individuals (14%) report a disability more than four times in 10 years, which is defined as a chronic disability by Meyer and Mok (2013).

5.2 Results

First, we consider the case that does not control for health status, in which health shocks are assumed to be non-separable from total income shocks and uncorrelated with individual-specific income components. This assumption is required not only for random effects estimation, but also for the fixed effects estimation when health status is not directly controlled for. This is because our individual-by-individual OLS estimations for fixed effects also require total income shocks, including health shocks, to be random.

Panel (1) of Table 4 shows the estimates obtained without extracting health shocks from total income shocks. The estimates of $\rho$ are 0.824, 0.798, and 0.831 for all individuals, the college educated, and the high school educated, respectively. The persistence of income shock is greater for high school-educated individuals, while their variance in long-lasting income shock is lower (0.037 compared to 0.043 for college educated individuals) and their temporary shocks vary more (0.048 compared to 0.042 for college educated individuals).\footnote{All of these estimates are very similar to Guvenen’s estimates, assuming random effects for the trend compo-}
Next we extract health shocks from total income shocks to investigate to what extent health shocks account for the persistence and variance of income shocks. We use fixed effects estimations to extract the disability variable together with individual-specific income components. Then, the MD estimation is carried out based on the autocovariances of the unknown components after removing the component of the impact of health shocks, $m$, as well as the individual components $f$. Note that the estimates are interpreted as the persistence and variance of the income shocks from which health shocks have been removed. Alternatively, one can also interpret the estimates for income shocks for the case in which no individual experiences any health shock.

Panel (2) of Table 4 shows the results. First, health shock reduces labor income by 5.7 percentage points for all individuals and by 5.1 and 5.9 percentage points for the college educated and high school educated, respectively. Labor earnings for the college educated seem to be better insured against health shocks compared to the high school educated, although we do not really know about the details because we do not consider the severity of the disability.

The magnitude of the effect may be lower than expected, especially if we consider the fact that the disability variable is defined as having a work limitation that could lower labor earnings. This may be because our sample contains only individuals who have worked for at least 20 years of the 29 years of the total sample period (not necessarily consecutive). The individuals in our sample may be a particular group of people who choose to remain in the labor force, despite their work limitation. The effect of health shocks can be underestimated due to the selection effect. Correcting such a sample selection is difficult, especially with heterogeneous income growth components, and beyond the scope of this paper.

We compare our main estimates from the MD estimation controlling for health status and allowing for fixed effects (panel (2) of Table 4) with the estimates without extracting health shocks.

---

26 We do not report the individual income growth rates, $\beta_i$, because these are biased for short $T$.
27 The measure of disability based on the survey question does not necessarily restrict respondents to drop out of the labor force.
28 They may or may not be eligible to apply for disability insurance but disability insurance requires a five-month waiting period between application and the receipt of benefits, during which period the individual must remain unemployed.
29 Low and Pistaferri (2012) do not allow for individual fixed effects, while they consider the sample selection to estimate the variances of income shocks controlling disability. A series of papers develops a bias correction method for sample selection with time-invariant fixed effects (e.g., Fernandez-Val and Vella, 2011; Semykina and Wooldridge, 2013). However, these papers do not consider individual-specific trends. Our case is more complicated because our main interests lie in the persistence and variance of shocks, so that MD estimation is additionally used.
from total income shocks (panel (1) of Table 4). In the strict sense, the estimates in panels (1) and (2) of Table 4 are not exactly comparable because the estimates in panel (1) can be biased due to possible correlation between health shocks and individual-specific income growth components. Nevertheless, the comparison can be useful because the literature usually presents the estimates in panel (1) under the assumption that non-separable income shocks (including health shocks) are purely random.

The estimates of $\rho$ decrease from 0.824 to 0.782 for all individuals, from 0.798 to 0.735 for the college educated, and from 0.831 to 0.803 when we extract health shocks from income shocks. These results suggest that health shocks contribute to the persistence of total income shock if they are included. However, in terms of the variance of permanent and transitory shocks, no difference is noticeable. This result may be surprising, but, theoretically, controlling for health shocks can have no impact on the variance, even though it has an impact on persistence. Again, these differences may be underestimated by the sample selection as well. In addition, classical measurement error, if any, in the variable of health would be captured in the transitory shock. This could then lead to an upward bias in the estimate of $\sigma_\varepsilon^2$, such that the impact can be underestimated as well.

Based on our estimates of $\rho$, Figure 1 illustrates the impact of permanent shocks as time goes by after their onset. The decay of the impact of an income shock to one-fifth of the initial impact takes eight years for college-educated individuals from its onset, without controlling for health shocks, while it takes 10 years for high school-educated individuals. Health shocks contribute up to two years to the persistence of income shocks for the college educated; income shocks from which health shocks are extracted decay to one-fifth of the initial impact in six years. Health shocks have less of an impact on the persistence of income shocks for high school-educated individuals, for whom it also takes eight and a half years for the impact of income shocks from which health shocks are extracted to drop to one-fifth of the initial impact.

6 Conclusion

In this paper, we provide an alternative econometrics methodology to estimate a standard HIP model allowing for fixed HIP coefficients. The proposed estimation method allowing for fixed effects has a methodological contribution because it is more general and less restrictive than the random
effects approach. Moreover, importantly, this method can be used to understand the sources of income shocks and their contribution to income fluctuations using the HIP model.

Recent developments in the studies of income dynamics attempt to understand the sources of income shocks, but these studies are mainly developed in the setting of restricted income profiles (RIP) models in which individuals face perfectly permanent income shocks (i.e., $\rho = 1$). The studies do not allow for changes in the level of income shock persistence when particular sources of income shocks are disentangled. Our methodology can be an alternative way to study the sources of income shocks allowing for HIP and changes in the degree of persistence of income shocks when we disentangle shocks. Although income shocks appear to be very persistent before disentanglement, such persistence may not hold once we remove particular sources of income shocks.

We apply our methodology to measure the variability and persistence of income shocks by explicitly controlling for individuals’ disability status as health shocks in the HIP model. Health shocks are an often-cited driver of permanent income shocks and, by including these in the income specification, we analyze to what extent health shocks account for the persistence of income shocks. Our estimation results suggest that health shocks in the form of disability indeed account for some of the persistence of income shocks. This application highlights the potential of the proposed fixed effect HIP model and may encourage others using a similar approach to identify the contribution of other sources of income risk.

Our application has limitations but extensions may be possible. Our sample is restricted to individuals who choose to remain in the labor force despite their work limitation. Such a sample selection might be considered in the future research. Our analysis is also based on the assumption that health shocks are exogenous after allowing for unknown individual-specific income components to be correlated with health shocks. This may still be a strong assumption, because various sources of income risk can lead to health status becoming endogenous to income. Other sources of income risk could be controlled together with health shocks, since our methodology does not limit the number of controls.
References


Table 1. Simulation: Fixed Effects MD estimation with Sampling Bias Correction

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho$</td>
<td>$\sigma^2_\eta$</td>
</tr>
<tr>
<td>True</td>
<td>0.750</td>
<td>0.050</td>
</tr>
<tr>
<td>Bias</td>
<td>0.005</td>
<td>0.000</td>
</tr>
<tr>
<td>Std</td>
<td>0.011</td>
<td>0.001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.012</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.850</td>
<td>0.050</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>Std</td>
<td>0.012</td>
<td>0.001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.013</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Two different data generating processes are considered: (1) with $x_{it}^*$ and (2) without $x_{it}^*$. 
Table 2. Simulation: Fixed Effects MD estimation without Sampling Bias Correction

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho$</td>
<td>$\sigma^2_{\eta}$</td>
<td>$\sigma^2_{\varepsilon}$</td>
</tr>
<tr>
<td>True</td>
<td>0.750</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.572</td>
<td>0.038</td>
<td>-0.042</td>
</tr>
<tr>
<td>Std</td>
<td>0.051</td>
<td>0.026</td>
<td>0.027</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.574</td>
<td>0.046</td>
<td>0.050</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho$</td>
<td>$\sigma^2_{\eta}$</td>
<td>$\sigma^2_{\varepsilon}$</td>
</tr>
<tr>
<td>True</td>
<td>0.850</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.635</td>
<td>0.042</td>
<td>-0.046</td>
</tr>
<tr>
<td>Std</td>
<td>0.044</td>
<td>0.019</td>
<td>0.020</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.636</td>
<td>0.046</td>
<td>0.050</td>
</tr>
</tbody>
</table>

(1) PSID replicated samples, $T=29$. $f$ is a quadratic function in experience.
(2) One cohort for $T=40$ without attrition. $f$ is a linear function in experience.
(3) One cohort for $T=100$ without attrition. $f$ is linear function in experience.

We minimize the objective function with the constraint for the estimated variances to be positive for the simulations in (1). For (2) and (3) the variances are estimated to be positive without the constraint.
Table 3. Disability Rates and Prevalence

A. Disability Rates (pooled years)

<table>
<thead>
<tr>
<th>Age</th>
<th>All</th>
<th>College</th>
<th>High School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 35</td>
<td>5.6</td>
<td>4.6</td>
<td>6.0</td>
</tr>
<tr>
<td>35-55</td>
<td>8.2</td>
<td>6.9</td>
<td>8.7</td>
</tr>
<tr>
<td>Above 55</td>
<td>13.0</td>
<td>13.1</td>
<td>12.9</td>
</tr>
<tr>
<td>All</td>
<td>7.6</td>
<td>6.6</td>
<td>8.1</td>
</tr>
</tbody>
</table>

B. Disability Prevalence (across individuals)

<table>
<thead>
<tr>
<th>Frequency</th>
<th>All</th>
<th>College</th>
<th>High School</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least once</td>
<td>39.7</td>
<td>33.7</td>
<td>42.4</td>
</tr>
<tr>
<td>One to three times</td>
<td>23.3</td>
<td>20.2</td>
<td>24.8</td>
</tr>
<tr>
<td>More than four times</td>
<td>16.4</td>
<td>13.5</td>
<td>17.6</td>
</tr>
<tr>
<td>More than four times for 10 years</td>
<td>14.1</td>
<td>10.8</td>
<td>15.6</td>
</tr>
</tbody>
</table>
Table 4. HIP MD Estimates for the models with and without disability variable

<table>
<thead>
<tr>
<th></th>
<th>Fixed Effect</th>
<th>Fixed Effect with Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>$\rho$</td>
<td>$\sigma^2_\eta$</td>
</tr>
<tr>
<td>All</td>
<td>0.824</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>College</td>
<td>0.798</td>
<td>0.043</td>
</tr>
<tr>
<td>Educated</td>
<td>(0.031)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>High school</td>
<td>0.831</td>
<td>0.037</td>
</tr>
<tr>
<td>Educated</td>
<td>(0.022)</td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

The individual income components are specified as a quadratic in experience.

Bootstrap Standard errors are in parentheses (with 100 repetitions).
Figure 1. The Decay of Income Shock
Appendix: Technical Derivations

Derivation of $\gamma (\theta, t, k, \tau)$

First, we derive the probability limit of $\hat{\gamma} (t, k, \tau)$ which is denoted by $\gamma (\theta, t, k, \tau)$. Recall that if individual $i$ belongs to cohort $\tau$ so that $h_{i\tau} = 1$, then the experience at time $t$ is

$$h_{it} = h_{t, \tau} = (t - \tau + 1) I \{ t \geq \tau \}$$

and

$$H_{it}^K = H^K (h_{t, \tau}) s_{it}.$$  

Notice that when $\{ x_{it}^* \}$ is strictly exogenous with respect to $\{ \xi_{it} \}$ and the missing observations are random, the fixed effect estimator $\hat{\pi}$ is consistent for $\pi$ and $N \to \infty$. Then, it also follows that $\hat{\beta}_i = \hat{\beta}_i (\pi) + o_p (1)$ as $N \to \infty$. From this, we can deduce that for all $i \in I (\tau)$ and $t \in T (\tau)$ we have

$$\hat{\xi}_{it} = \xi_{it} - H_{it}^{K_f} \left( \sum_{t \in T (\tau)} H_{it}^K H_{it}^{K_f} \right)^{-1} \sum_{s \in T (\tau)} H_{is}^K \xi_{is}.$$  

Therefore, for individuals of cohort $\tau$,

$$\hat{\xi}_{it} = s_{it} p_{it}^* - s_{it} H_{t, \tau} \left( \sum_{t \in T (\tau)} H_{t, \tau} H_{t, \tau}^t s_{it} \right)^{-1} \sum_{s \in T (\tau)} H_{s, \tau} p_{is}^* s_{is}$$

$$+ s_{it} e_{it}^* - s_{it} H_{t, \tau} \left( \sum_{t \in T (\tau)} H_{t, \tau} H_{t, \tau}^t s_{it} \right)^{-1} \sum_{s \in T (\tau)} H_{s, \tau} e_{is}^* s_{is}$$

$$= \hat{p}_{it} + \hat{e}_{it}, \text{ say.}$$

Notice that $\gamma (\theta, t, k, \tau) = \mathbb{E} (\hat{\xi}_{it} | h_{i\tau} = 1) = \mathbb{E} (\hat{p}_{it} \hat{p}_{it-1} | h_{i\tau} = 1) + \mathbb{E} (\hat{e}_{it} \hat{e}_{it-1} | h_{i\tau} = 1)$. In what follows we derive $\mathbb{E} (\hat{p}_{it} \hat{p}_{it-1} | h_{i\tau} = 1)$ and $\mathbb{E} (\hat{e}_{it} \hat{e}_{it-1} | h_{i\tau} = 1)$.

For each $i$ of cohort $\tau$, define

$$f_{i, T} (t, s; \tau) = s_{it} H_{i, \tau} \left( \sum_{t \in T (\tau)} H_{i, \tau} H_{i, \tau}^t s_{it} \right)^{-1} H_{s, \tau} s_{is}.$$  

Define $P_{\tau}^{s_1, \ldots, s_l}$ to be the probability that the samples of time period $s_1, \ldots, s_l$ are observed when the individual belongs to cohort $\tau$,

$$P_{\tau}^{s_1, \ldots, s_l} = \mathbb{P} (s_{is_1} = \cdots = s_{is_l} = 1 | h_{i\tau} = 1).$$
Notice that for individual \(i\) of cohort \(\tau\), we have

\[
\mathbb{E}(\tilde{p}_{it} \tilde{p}_{it-1} \mid h_{i\tau} = 1)
= \mathbb{E}(p_{it} p_{it-1} \mid s_{it} = s_{it-1} = 1, h_{i\tau} = 1) \mathbb{P}(s_{it} = s_{it-1} = 1 \mid h_{i\tau} = 1)
- \sum_{s \in T(c)} \mathbb{E}(p_{it} p_{is} f_{i,T}(t - k, s; \tau) \mid s_{it} = s_{it-1} = s_{is} = 1, h_{i\tau} = 1) \mathbb{P}(s_{it} = s_{it-1} = s_{is} = 1 \mid h_{i\tau} = 1)
- \sum_{s \in T(c)} \mathbb{E}(p_{it-k} p_{is} f_{i,T}(t, s; \tau) \mid s_{it} = s_{it-k} = s_{is} = 1, h_{i\tau} = 1) \mathbb{P}(s_{it} = s_{it-k} = s_{is} = 1 \mid h_{i\tau} = 1)
+ \sum_{s \in T(c)} \sum_{w \in T(c)} \left[ \mathbb{E}(f_{i,T}(t, s; \tau) f_{i,T}(t - k, w; \tau) p_{is} p_{iw} \mid s_{it} = s_{it-k} = s_{is} = s_{iw} = 1, h_{i\tau} = 1) \right] \mathbb{P}(s_{it} = s_{it-k} = s_{is} = s_{iw} = 1 \mid h_{i\tau} = 1)
\]

where the last equality holds since \(p_{it} p_{it-1}\) is independent of \(f_{i,T}(t_3, t_4, \tau)\) conditional on \((s_{t_1}, s_{t_2}, h_{i\tau} = 1)\).

Denote

\[
\gamma_p (\rho, t, s; \tau) = \frac{1}{\sigma^2} \mathbb{E}(p_{it}^* p_{is}^*) ,
\]
a standardized covariance of the long lasting shocks of time \(t\) and \(s\) of individuals of cohort \(\tau\). Recall that for individual \(i\) of cohort \(\tau\), we have

\[
p_{it}^* = p_{it}^* (h_{t,\tau}) = \eta_{it}^* (h_{t,\tau}) + \rho \eta_{it-1}^* (h_{t,\tau} - 1) + \ldots + \rho^{h_{t,\tau} - 1} \eta_{i\tau}^* (1)
\]

for \(t \in T(\tau)\). Then,

\[
\gamma_p (\rho, t, s; \tau) = \rho^{t-s} \left( 1 + \rho^2 + \ldots + \rho^{2(h_{\min(s,t),\tau} - 1)} \right) \mathbb{1}\{h_{\min(s,t),\tau} \geq 1\}
= \rho^{t-s} \frac{(1 - \rho^{2h_{\min(s,t),\tau}})}{1 - \rho^2} \mathbb{1}\{h_{\min(s,t),\tau} \geq 1\}.
\]
Notice that for $i$ with cohort $\tau$, we have

$$
\mathbb{E}(p_{it}p_{it-k}|s_{it} = s_{it-k} = 1, h_{i\tau} = 1) = \mathbb{E}(p_{it}^*p_{it-k}^*|s_{it} = s_{it-k} = 1, h_{i\tau} = 1)
= \mathbb{E}(p_{it-k}^*)
= \sigma^2_p \gamma_p (\rho, t, t-k; \tau)
$$

$$
\mathbb{E}(p_{it}p_{is}|s_{it} = s_{it-k} = s_{is} = 1, h_{i\tau} = 1) = \mathbb{E}(p_{it}^*p_{is}^*|s_{it} = s_{it-k} = s_{is} = 1, h_{i\tau} = 1)
= \mathbb{E}(p_{it}^*)
= \sigma^2_p \gamma_p (\rho, t, s; \tau)
$$

$$
\mathbb{E}(p_{it-k}p_{is}|s_{it} = s_{it-k} = s_{is} = 1, h_{i\tau} = 1) = \mathbb{E}(p_{it-k}^*p_{is}^*|s_{it} = s_{it-k} = s_{is} = 1, h_{i\tau} = 1)
= \mathbb{E}(p_{it-k}^*)
= \sigma^2_p \gamma_p (\rho, t-k, s; \tau)
$$

$$
\mathbb{E}(p_{is}p_{iw}|s_{it} = s_{it-k} = s_{is} = s_{iw} = 1, h_{i\tau} = 1) = \mathbb{E}(p_{is}^*p_{iw}^*|s_{it} = s_{it-k} = s_{is} = s_{iw} = 1, h_{i\tau} = 1)
= \mathbb{E}(p_{is}^*)
= \sigma^2_p \gamma_p (\rho, s, w; \tau).
$$

Therefore, we have for individual $i$ of cohort $\tau$, we have

$$
\mathbb{E}(\tilde{p}_{it}\tilde{p}_{it-k}|h_{i\tau} = 1) = \sigma^2_p \gamma_p (\rho, t, t-k; \tau) \mathbf{P}_{\tau}^{t,t-k}
- \sigma^2_p \sum_{s \in \mathcal{T}(c)} \gamma_p (\rho, t, s; \tau) \mathbb{E}(f_{i,T}(t-k, s, \tau)|s_{it} = s_{it-k} = s_{is} = 1, h_{i\tau} = 1) \mathbf{P}_{\tau}^{t,t-k,s}
- \sigma^2_p \sum_{s \in \mathcal{T}(c)} \gamma_p (\rho, t-k, s; \tau) \mathbb{E}(f_{i,T}(t, s, \tau)|s_{it} = s_{it-k} = s_{is} = 1, h_{i\tau} = 1) \mathbf{P}_{\tau}^{t,t-k,s}
+ \sigma^2_p \sum_{s \in \mathcal{T}(c)} \sum_{w \in \mathcal{T}(c)} \left[ \gamma_p (\rho, s, w; \tau) \mathbb{E}(f_{i,T}(t, s, \tau) f_{i,T}(t-k, w, \tau)|s_{it} = s_{it-k} = s_{is} = s_{iw} = 1, h_{i\tau} = 1) \right] \mathbf{P}_{\tau}^{t,t-k,s,w}.
$$
Similarly, we have

\[
\mathbb{E}(\hat{e}_{it\bar{k}}\mid h_{it\tau} = 1) = \mathbb{E}(e_{it\bar{k}}\mid s_{it\bar{k}} = s_{it} = 1, h_{it\tau} = 1) \mathbb{P}(s_{it} = s_{it\bar{k}} = 1 \mid h_{it\tau} = 1) - \sum_{s \in T(c)} \mathbb{E}(e_{it\bar{k}}f_{i,T}(t - k, s; \tau) \mid s_{it\bar{k}} = s_{is} = 1, h_{it\tau} = 1) \mathbb{P}(s_{it\bar{k}} = s_{is} = 1 \mid h_{it\tau} = 1) - \sum_{s \in T(c)} \mathbb{E}(e_{it\bar{k}}\mid s_{it\bar{k}} = s_{is} = 1, h_{it\tau} = 1) \mathbb{P}(s_{it\bar{k}} = s_{is} = 1 \mid h_{it\tau} = 1) + \sum_{s \in T(c)} \sum_{w \in T(c)} \left[ \mathbb{E}(f_{i,T}(t, s; \tau) f_{i,T}(t - k, w; \tau) e_{is}e_{iw} \mid s_{it\bar{k}} = s_{is} = s_{iw} = 1, h_{it\tau} = 1) \times \mathbb{P}(s_{it\bar{k}} = s_{is} = s_{iw} = 1 \mid h_{it\tau} = 1) \right].
\]

Denote

\[
\gamma_c(\mu, t, s; \tau) = \frac{1}{\sigma_c^2} \mathbb{E}(e_{it\bar{k}}e_{is\bar{k}}^*),
\]

a standardized covariance of the temporal shocks at time \(t\) and \(s\) of cohort \(\tau\). Recall that for individual \(i\) of cohort \(\tau\),

\[
e_{it\bar{k}} = e_{it\bar{k}}(h_{it\tau}) = e_{it\bar{k}}(h_{it\tau}) + \phi e_{it\bar{k} - 1}(h_{it\tau} - 1)
\]

with \(e_{it\bar{k} - 1}(0) = 0\). Then,

\[
\gamma_c(\phi, t, s; \tau) = \begin{cases} 
1 & \text{if } t = s = \tau \\
1 + \phi^2 & \text{if } t = s, t > \tau \\
\phi & \text{if } |t - s| = 1, \min\{t, s\} \geq \tau
\end{cases}
\]

Also, we have

\[
\mathbb{E}(e_{it\bar{k}}\mid s_{it\bar{k}} = s_{it\bar{k}} = 1, h_{it\tau} = 1) = \mathbb{E}(e_{it\bar{k}}\mid s_{it\bar{k}} = s_{it\bar{k}} = 1, h_{it\tau} = 1) = \mathbb{E}(e_{it\bar{k}}e_{it\bar{k}}^*) = \sigma_c^2 \gamma_c(\phi, t, t - k; \tau)
\]

34
\[ \mathbb{E}(e_{it}e_{is}|s_{it} = s_{it-k} = s_{is} = 1, h_{ir} = 1) 
= \mathbb{E}(e_{it}^*e_{is}^*|s_{it} = s_{it-k} = s_{is} = 1, h_{ir} = 1) 
= \mathbb{E}(e_{it}^*) 
= \sigma^2_e \gamma_e(\phi, t, s; \tau) \]

\[ \mathbb{E}(e_{it-k}e_{is}|s_{it} = s_{it-k} = s_{is} = 1, h_{ir} = 1) 
= \mathbb{E}(e_{it-k}^*e_{is}^*|s_{it} = s_{it-k} = s_{is} = 1, h_{ir} = 1) 
= \mathbb{E}(e_{it-k}^*) 
= \sigma^2_e \gamma_e(\phi, t - k, s; \tau) \]

\[ \mathbb{E}(e_{is}e_{iw}|s_{it} = s_{it-k} = s_{is} = s_{iw} = 1, h_{ir} = 1) 
= \mathbb{E}(e_{is}^*e_{iw}^*|s_{it} = s_{it-k} = s_{is} = s_{iw} = 1, h_{ir} = 1) 
= \mathbb{E}(e_{is}^*) 
= \sigma^2_e \gamma_e(\phi, s, w; \tau). \]

Therefore, for individual \( i \) of cohort \( \tau \),

\[ \mathbb{E}(\tilde{e}_{it}\tilde{e}_{it-k}|h_{ir} = 1) \]

\[ = \sigma^2_e \gamma_e(\phi, t, t - k; \tau) P_{\tau}^{t,t-k} 
- \sigma^2_e \sum_{s \in T(c)} \gamma_e(\phi, t, s; \tau) \mathbb{E}(f_{it,T}(t - k, s; \tau) | s_{it} = s_{it-k} = s_{is} = 1, h_{ir} = 1) P_{\tau}^{t,t-k,s} 
- \sigma^2_e \sum_{s \in T(c)} \gamma_e(\phi, t - k, s; \tau) \mathbb{E}(f_{it,T}(t, s; \tau) | s_{it} = s_{it-k} = s_{is} = 1, h_{ir} = 1) P_{\tau}^{t,t-k,s} 
+ \sigma^2_e \sum_{s \in T(c)} \sum_{w \in T(c)} \left[ \gamma_e(\phi, s, w; \tau) \mathbb{E}(f_{it,T}(t, s; \tau) f_{i,T}(t - k, w; \tau) | s_{it} = s_{it-k} = s_{is} = s_{iw} = 1, h_{ir} = 1) \times P_{\tau}^{t,t-k,s,w} \right]. \]

This gives

\[ \gamma(\theta, t, k, \tau) = \mathbb{E}(\tilde{p}_{it}\tilde{p}_{it-k}|h_{ir} = 1) + \mathbb{E}(\tilde{e}_{it}\tilde{e}_{it-k}|h_{ir} = 1), \]

where \( \mathbb{E}(\tilde{p}_{it}\tilde{p}_{it-k}|h_{ir} = 1) \) is (7) and \( \mathbb{E}(\tilde{e}_{it}\tilde{e}_{it-k}|h_{ir} = 1) \) in (8).

**Computation of** \( \hat{\gamma}(\theta, t, k, \tau) \)

Notice that the probability limit \( \gamma(\theta, t, k, \tau) \) contains unknown components such as

\[ \mathbb{E}(f_{i,T}(t, s; \tau) | s_{it} = s_{is} = s_{iw} = 1, h_{ir} = 1) \]

and \( P_{\tau}^{t,s,w} \). The \( \hat{\gamma}(\theta, t, k, \tau) \) replaces the unknown components with their estimators.

Recall that \( N(\tau) = \sum_{i=1}^{N} 1 \{ h_{ir} = 1 \} \), the number of individuals of cohort \( \tau \). Let \( \mathcal{I}_{t_1,\ldots,t_k}(\tau) = \{ i \in \mathcal{I}(\tau) : s_{it_1} = \cdots = s_{it_k} = 1, h_{ir} = 1 \} \) be the collection of the individuals who belong to cohort \( \tau \) and whose \( t_1,\ldots,t_k \) \( \in \mathcal{I}(\tau) \) time period samples are observed and \( N_{t_1,\ldots,t_k}(\tau) = \sum_{i \in N(\tau)} 1 \{ s_{it_1} = \cdots = s_{it_k} = 1, h_{ir} = 1 \} \) be the number of the individuals of cohort \( \tau \) such that time periods \( \{ t_1,\ldots,t_k \} \) are all observed.
As estimators of the unknown components $P_{\tau}^{t,t-k,s}$ and $P_{\tau}^{t,t-k,s,w}$ we suggest

$$\hat{P}_{\tau}^{t,t-k,s} = \frac{N_{t,t-k,s}(\tau)}{N(\tau)}$$

$$\hat{P}_{\tau}^{t,t-k,s,w} = \frac{N_{t,t-k,s,w}(\tau)}{N(\tau)},$$