The Evolution of Canadian Income Dynamics: Evidence from Canadian Longitudinal Tax Records*

Brant Abbott and Shi Chen

Preliminary Draft - October 17, 2025

{Access Latest Version Here}

Abstract

What do income dynamics look like for Canadian-born workers since the 1980s? We address this question by estimating a flexible, nonlinear income model using panel data constructed from Canadian administrative tax records. A novel feature of our approach is the introduction of a proxy for Canadian post-secondary education attainment, imputed using records of education deductions and tax credits. We find that pre-tax household labour income exhibits strong state- and history-dependent persistence. Specifically, high-rank income shocks have long-lasting effects only for high-income households, while low-rank income shocks decay over time at heterogenous rates that depend on current income states. The non-permanence of adverse income shocks appears to be unique to Canadian working families and remains remarkably stable over the period 1985 to 2019. Tracking workers from the 1964 birth cohort across their life-cycle, we find that the growth of income persistence occurs primarily within the first decade of their working lives, a pattern that is consistent across subsequent cohorts. Finally, analysis of cohorts that experienced major recessions during their working years reveals moderate changes in income persistence during those recessionary periods.

*This study uses micro-level administrative dataset provided by Statistics Canada. We thank the Queen's Research Data Centre (QRDC) and the Canadian Research Data Centre Network (CRDCN) for their excellent assistance on data access. The views expressed herein are those of the authors and do not reflect the views of Statistics Canada or CRDCN. All errors are our own.

1 Introduction

Income fluctuations are the primary source of idiosyncratic risk for workers. Individuals who rely on labour income are inherently vulnerable to income shocks, and differences in the persistence and magnitude of these shocks lead to cross-sectional income dispersions and long-term changes in income inequality. The design of policy programs intended to mitigate the effects of unforeseen income shocks relies on a correct representation of the income process to accurately reflect the extent of risks faced by working individuals.

In this paper, we study the time evolution of income persistence for Canadian-born workers using administrative tax records. We focus on two dimensions of time: chronological time and age. First, we compare income persistence across five groups of Canadian-born workers sampled over a 35-year window using rolling panels. This approach allows us to identify both the common properties of the income process shared by workers over time, as well as subtle features unique to specific time periods such as recessions. Second, we follow the 1964 birth cohort over most of its working life attempt to examine how income levels and shock persistence evolves as workers age and advance in their career path. To evaluate how representative this cohort is, we also analyze cohorts born in 1971, 1978 and 1985 (spaced seven years apart to avoid overlapping panels) to isolate features of income persistence that consistently appear at specific ages. By focusing on pre-tax labour income, we abstract from changes arising from tax reforms and ensure that labour income is measured consistently throughout the entire sample period.

Our empirical approach enables us to explore multiple dimensions of heterogeneity in the income dynamics. We model labour income as a general first-order Markov process following Arellano, Blundell, and Bonhomme (2017). This flexible model specification allows income persistence to vary with an individual's income history, as well as the rank of current income shocks. We use two complementary measures of income to capture different aspects of income evolution. The first measure is the residual labour income in nature logarithm, net of the deterministic component explained by observable changes in worker and household characteristics. The second measure further decomposes residual log income into persistent and transitory components, with the former being the main driver of serial correlation in residual labour income. Directly comparing results based on these two measures sheds light on the extent to which transitory shocks affect income persistence.

A key requirement for studying income dynamics is access to panel data with consistent measure of income over time. Traditional Canadian data sources are limited in this regard: they are either cross-sectional (such as the Labour Force Survey, "LFS"), collected at a lower frequently (such as the Census at five-year frequency), or have been discontinued with limited availability (such as the Survey of Labour and Income Dynamics, "SILD"). Our study instead uses the Longitudinal Administrative Databank (LAD) provided by Statistics Canada. The LAD is a 20% random sample of all Canadian tax filers and contains detailed information from their personal income tax forms. Its exceptional panel length, spanning from 1982 to 2019, is second-to-none and generally unmatchable by any survey-based datasets. Moreover, the extent of measurement error and inconsistency in the LAD is expected to be minimal given the nature of tax records collected and managed by the Canadian tax authority.

One limitation of the LAD is the absence of direct information on education. To address this issue, we impute a proxy for Canadian post-secondary education attainments, which is crucial given the well-documented relationship between education levels and income. Our imputation strategy relies on historical information on education-related deductions and tuition tax credits reported by individual tax filers. Both full-time and part-time students attending Canadian post-secondary education institutions can claim education deductions based on the number of months enrolled and use their tuition fees as non-refundable tax credits. As a result, it is possible to reconstruct the total number of months of enrolments over a consecutive period and identify the fee types paid at the time of enrolment wherever historical tuition data are available. Combined with information on typical program duration, we can infer the type of post-secondary education received by an individual tax filer and whether the program was completed. Because the LAD follows individuals over extended periods, we can recover the highest level of post-secondary education attained in Canada with great accuracy by observing the entire sequence of education events.

We find that income dynamics for Canadian-born workers remain largely stable over the period of 1985-2019. Their income persistence exhibits substantial state and history dependence, meaning that income shocks of different ranks (i.e. percentiles of shock distribution) would affect working households differently depending on their income history from the previous tax year. Specifically, positive income shocks (with high ranks) have permanent effects only for high income households, whereas negative income shocks (with low ranks) do not display unit root behaviour. This is a surprising feature of Canadian income dynamics, as the permanence of bad income shocks for low income households has been documented in the U.S. and Norwegian data. Furthermore, we find that workers born in 1964 experienced a period of lower income persistence only during the first 10 years of their working life, after which income persistence remained relatively stable as they age. Nearly identical patterns are observed for workers born in subsequent decades with small cohort-to-cohort variations. The impact of recessionary periods is visible in our results, primarily represented by an increase in the persistence of negative income shocks for high income households. Given the large sample size available from the LAD, we find our empirical results are precisely estimated.

Our work contributes to the growing literature on Canadian income inequality, mobility and dynamics. Early studies examined various types of income measures: Blackburn and Bloom (1993) and Gottschalk (1993) focused on family income, Picot (2001) on income from Canadian men and women separately and Baker and Solon (2003) on Canadian men only. These studies find that earnings inequality in Canada rose in the 1980s and early 1990s, driven mainly by changes at the top of the earnings distribution. Richardson (1997) documented a clear upward trend in wage inequalities from 1981 to 1992 for Canada, while Fortin and Lemieux (2015) examined inter-provincial differences in wage growth and dispersion. In terms of micro-data usage, Morissette and Berube (1996) were the first paper to use Canadian tax records derived from T4 employment income forms and many subsequent papers have followed their idea of administrative datasets. More recently, Bowlus, Gouin-Bonenfant, Liu, Lochner, and Park (2022) studies the evolution of individual earnings inequality in Canada from 1983 to 2016. They find modest changes in income inequality and a strong correlation between income dynamics and the firm characteristics. The increasing availability of panel income data and advances of econometric methods promoted estimation of flexible income process. Arellano et al. (2017) provides the theoretical foundations for a nonlinear income process which we build upon here. They also provided evidences of state and history dependent income persistence identified in U.S. survey data and Norwegian administrative data. Kitao, Suzuki, and Yamada (2025) studies the nonlinearities in individual pre-tax incomes using Japanese municipal tax records. To our knowledge, we are the first to estimate nonlinear income dynamics using Canadian administrative data.

The remainder of this paper is organized as follows. Section 2 introduces our empirical framework of nonlinear income model. Section 3 discusses the construction of different samples based on Canadian tax records and describes our approach to create a proxy for post-secondary education. Section 4 presents results on nonlinear income persistence, separately for historical, life-cycle and cohort-based samples. Section 5 concludes.

2 Empirical methodology

Our empirical methodology is based on the nonlinear income model proposed by Arellano et al. (2017). We consider an income measure for individual i at time t, denoted $z_{i,t}$, with the following first-order Markov structure:

$$z_{i,t} = g(z_{i,t-1}, u_{i,t}) \tag{1}$$

where $u_{i,t}$ is an income shock normalized to $\mathcal{U}(0,1)$, independent of the entire history of past shocks $u^{t-1} = (u_{i,1}, \ldots, u_{i,t-1})$ as well as future shocks. $g(.)$ is a nonlinear function used to capture both the history-dependence of z and the impact of income shocks to z. A common form of g used to study income dynamics follows the first-order autoregressive (AR(1)) process:

$$z_{i,t} = \tilde{\rho}(z_{i,t-1}, u_{i,t}) z_{i,t-1} + u_{i,t} \tag{2}$$

such that the income persistence $\tilde{\rho}$ is both history dependent (through $z_{i,t-1}$) and state dependent (through $u_{i,t}$). Note that the income model described in Eq 1 is more general than Eq 2 and the autoregressive form is not required or enforced in the estimation.

We are interested in the average persistence of a given income measure. Following the model described in Eq 1, the income persistence of z can be computed based on the following partial derivative of g:

$$\rho(z_{i,t-1}, u_{i,t}) = \frac{\partial g(z_{i,t-1}, u_{i,t})}{\partial z_{i,t-1}} \tag{3}$$

where the persistence term ρ is nonlinear it its arguments and heterogenous in two dimensions

($z_{i,t-1}$ and $u_{i,t}$). This feature allows us to see how identical income shocks affect individual workers with distinct income histories differently over time. Another advantage of this nonlinear model is that it can capture the dynamic effects of income shocks. For the purpose of illustration, consider the nonlinear income model represented using the AR(1) form and repeat the substitution for $z_{i,t-1}$:

$$z_{i,t} = \rho(z_{i,t-1}, u_{i,t})z_{i,t-1} + u_{i,t} \tag{4}$$

$$= \rho(z_{i,t-1}, u_{i,t})\rho(z_{i,t-2}, u_{i,t-1})z_{i,t-2} + \rho(z_{i,t-1}, u_{i,t})u_{i,t-1} + u_{i,t} \tag{5}$$

Clearly, the contemporaneous effect of previous income shocks $u_{i,t-1}$ depends on current income persistence $\rho(z_{i,t-1}, u_{i,t})$, which in turn depends on current income shocks $u_{i,t}$. This implies that the effect of a historical income event can vary substantially depending on income shocks realized today. For example, a low income individual receiving a good income shock (such as getting a better job offer or promotion) today would not only receive a higher current income, she will also see a weakened effects of past negative income shocks as she begins to climb the job ladder. Another example would be health shocks. An individual receiving adverse health shock today may be force to take a medical leave, such that she would experience a loss of income today, as well as the declining benefits from positive income shocks occurred in the past as she progressed through her career. In alternative models where income persistence is the same for all individuals, there are no heterogenous effects of current income shocks and the dynamic effect of past income shocks are restricted to follow a fixed decay rate, which we find overly restrictive.

We use two types of income measures in our study. Let $Y_{i,t}$ denote the observed pre-tax labour income in natural logarithm and $X_{i,t}$ denote a vector of observed individual characteristics, we model $Y_{i,t}$ as:

$$Y_{i,t} = X_{i,t}\beta + y_{i,t} \tag{6}$$

where $X_{i,t}\beta$ captures the deterministic part of labour income that can be explain by individual characteristics (such as age, sex, education). This residual log income $y_{i,t}$ is our first income measure as it captures the stochastic part of labour income driven by shocks. Based on Eq 1, we estimate an income process as:

$$y_{i,t} = g^y(y_{i,t-1}, u_{i,t}^y) \tag{7}$$

where g^y is a non-parametric function. Our second income measure is based on a further distinction between two types of income shocks. We distinguish income shocks with long-lasting effects from those with only contemporaneous effects. The former type is referred to as the "persistent" income shock and the later type is recognized as "transitory" income shocks. We model $y_{i,t}$ as a sum of two components:

$$y_{i,t} = \eta_{i,t} + \varepsilon_{i,t} \tag{8}$$

$$\eta_{i,t} = g^\eta(\eta_{i,t-1}, u_{i,t-1}^\eta) \tag{9}$$

where η is the persistent income component with persistent income shock u^η. $\varepsilon_{i,t}$ is the transitory income component and equals to the transitory shock by construction. We assume the transitory component ε is i.i.d., serially uncorrelated and independent to persistent income shocks. η is our second income measure since the auto-correlation of residual labour income over time is entirely driven by η (through the non-parametric function g^η). We estimate the income model as characterized in Eq 9. Note that both η and ε are latent variables that are not directly observed by economists. We address this issue in later part of this section.

Following Arellano, Blundell, Bonhomme, and Light (2024), we adapt a quantile representation for both income models. We rewrite Eq 7 and Eq 9 as quantile functions of corresponding income measures:

$$y_{i,t} = Q^y(y_{i,t-1}, \tau), \tau \sim \mathcal{U}(0,1) \tag{10}$$

$$\eta_{i,t} = Q^\eta(\eta_{i,t-1}, \tau), \tau \sim \mathcal{U}(0,1) \tag{11}$$

where the income shocks u^y and u^η are normalized without loss of generality as standard uniform to conform with the quantile representation. Both non-parametric quantile functions Q^y and Q^η are approximated using lower-order Hermite polynomials.

Given panel data on income, the identification of Eq 10 is straightforward since the residual log incomes are observed in the data and Q^y is the conditional quantile function of $y_{i,t}$ given $y_{i,t-1}$. The identification of persistent income model (Eq 11) is established in Arellano et al. (2017) based on panel measurement error models.

Estimation of Eq 10 is straightforward since residual log income $y_{i,t}$ and $y_{i,t-1}$ are both observed. The income model represented by Q^y is estimated as a non-parametric quantile autoregression:

$$Q^y(y_{i,t-1},\tau) = \sum_{k=0}^{K} \alpha_k^y(\tau)\phi_k(y_{i,t-1}) \tag{12}$$

where ϕ_k denotes the k-th order Hermite polynomials. Income variables are standardized to mean zero and variance one. We specify the second income model from Eq 11 similarly:

$$Q^\eta(\eta_{i,t-1},\tau) = \sum_{k=0}^{K} \alpha_k^\eta(\tau)\phi_k(\eta_{i,t-1}, \text{age}_{i,t}) \tag{13}$$

where we allow for the quantile function of η taking worker's age as an additional argument to control for the life-cycle variations in income dynamics. In both models, we set τ to 11 different values to capture the heterogeneities in income dynamics between low and high income households.

Estimation of Eq 13 involves latent income variables $\{\eta, \varepsilon\}$ as controls and therefore requires the implementation of a stochastic version of the EM algorithm (sEM). This sEM algorithm features a two-step, iteratively approach. Consider the s-th iteration ($s \geq 1$):

0. (initialization) starting with a parameter vector $\hat{\alpha}^\eta_{(s)}$ for the income model in 13
1. (stochastic E-step) for every household i, draw M parallel copies of persistent income sequences $\eta_i^{(m)} = (\eta_{i,1}^{(m)}, \ldots, \eta_{i,T}^{(m)})$ from the posterior distribution $f_i(\eta_i^T | y_i^T, \text{age}_i^T; \hat{\alpha}^\eta_{(s)})$
2. (M-step) compute a new set of parameter estimates $\hat{\alpha}^\eta_{(s+1)}$ given the M parallel draws of $\eta_i^{(m)}$ by minimizing the quantile loss function

In the first step, the unobserved income variables η and ε are imputed using valid draws from their posterior distributions based on current model parameter estimates. In the second step, the imputed income components are used to update the quantile income models. New model parameter estimates from the current M-step is used in the E-step of next iteration to better approximate the posterior distribution. In practice, the first step is achieved using Particle Filter (Sequential Importance Resampling) and we iterate between both steps for a large number of times, while allowing for an extended burn-in period. Inference are based on non-parametric bootstrap of the entire algorithm and the unit of resampling is at the individual level (i.e. the entire sequence of observed income events for a particular individual). See Arellano et al. (2024) for more details on the sEM algorithm.

3 Data, proxy for education and sample creation

3.1 Longitudinal Administrative Databank (LAD)

We use individual tax records and self-reported demographic information from the Longitudinal Administrative Databank (LAD). LAD is a longitudinal dataset constructed by Statistics Canada from a subset of the Canadian T1 Family File (T1FF), which is compiled and maintained by the Canadian tax authority for the administration of income tax returns. According to the official documentation, LAD is a 20% random sample of the T1FF. Individual tax filers are uniquely identified and followed over time based on personal identifiers, providing an unparalleled advantage in terms of sample size, consistency and accuracy.

LAD contains detailed information on personal income, capital gains, transfers, as well as tax credits and deduction claimed. Given our focus on income dynamics, we concentrate on labour income aggregated at the parents/spouses level, which closely resembles the household labour income definition used in studies based on the Panel Study of Income Dynamics (PSID). We focus specifically on the pre-tax labour income, as Canadian tax legislation and the design of T1 form have undergone multiple rounds of reform since the LAD was first introduced. Focusing on pre-tax incomes rather than after-tax, disposable income helps reduce the influence of tax policy changes. The LAD also includes basic individual and family demographic variables such as age, sex, marital status, province of residence, etc., which we use as controls for observable factors influencing pre-tax labour income.

A key limitation of LAD is the absence of information on race and educational attainments, since tax filers are neither required nor able to report this information when filing for income tax returns. It is well established that the race and education significantly influence labour income through labour market discrimination and differences in human capital. There are evidences suggest that racial income gaps among Canadian-born workers remain visible (Pendakur and Pendakur (2007), Bonikowska et al. (2024)), although very few studies (e.g. Zhang (2014)) examine differences in income dynamics (as opposed to income levels) for Canadian-born visible minority workers. Due to the limited availability on personal information, it is generally infeasible to recover racial indicators from the LAD. We address the missing education information by creating a proxy of Canadian post-secondary education for individuals who have previously filed for education related tax credits and deductions.

3.2 Inferring Canadian post-secondary education from tax records

Despite the LAD's large sample size and long time horizon, the lack of education history remains a significant limitation of administrative datasets. Because educational attainment affects both income levels and trajectories, it is necessary to control for it in our empirical estimation. To do so, we extend the approach proposed by Finnie and Pavlic (2013) to infer a proxy for Canadian post-secondary education (PSE) based on tuition-related variables from individual tax records. Identifying education levels below post-secondary is extremely challenging with the information available in the LAD.

Our approach relies on the identification of two characteristics of PSE: the duration of study (measured in months enrolment) and the tuition fee types (relative to undergraduate tuition levels). For now, we restrict attention to full-time students. We begin with the variable `EDUDN` ("Educational deduction for full-time student"). According to documentation, "A full-time student at a designed education institution who is also enrolled in a qualifying educational program can claim particular amount of education deduction when filing their personal income tax returns". This amount is calculated as:

$$\texttt{EDUDN}_{i,t} = \texttt{Full\_time\_months}_{i,t} \times \texttt{Deduction\_per\_month}_t$$

where the `Deduction_per_month` varies across tax years and is available publicly. This allows us to recover the number of full-time enrolment months for each tax year and for each individual with positive education deductions. The number of months ranges from one month to twelve months in a given tax year. We exclude a small fraction of individuals who claimed more than thirteen months of full-time deductions or less than a full month equivalent. This procedure enables us to identify all episodes of PSE enrolments regardless of the actual duration of their studies. We refer to these enrolment months as "attendance" measures. Many individuals follow a common PSE attendance pattern: claiming tuition deductions for four consecutive years with eight months of full-time enrolment each year. Others display more complicated attendance patterns, with gaps years or switching of programs or institutions. We utilize all episodes of PSE attendance to infer their highest level of educational attainment, allowing for the possibility that an individual participated in multiple stages of PSE in Canada. Since the per-month deduction amount does not vary by the institution type (e.g. college vs. university), this variable alone cannot be used to recover the exact type of program in which an individual was enrolled in at the time of claim.

We turn to the amount of tuition paid to infer the type of institution or program an individual may have attended. The LAD contains a variable measuring tuition payments made by individuals in a given tax year. According to documentation, `TUTDN` ("Tuition fees for self") represents "A tuition fee is a non-refundable tax credit. If the tax filer was a student during the tax year, he or she may claim tuition fees (not books or expenses) paid to an educational institution of post-secondary level in Canada.". Therefore, individuals who attended a Canadian PSE institution and paid any tuition will have strong incentive to file these payments without understatement. Notice that these tuition payments are reported "as is", therefore reflect the number of months (or semesters) of enrolment at a PSE institution. To maintain consistency, we standardize these payments into eight-month equivalents by first dividing `TUTDN` by the number of full-time enrolment months, then multiplied by eight (as suggested by Finnie and Pavlic (2013)).

Tuition amounts by themselves are not informative as they need to be classified into different categories. We consider three categories: below-university (college, certificate or training programs), university and above-university. The university category would include two distinct levels of PSE: undergraduate (Bachelor's) and graduate (Master's or Doctoral) programs. Clearly, identifying the tuition ranges associated with these PSE categories would allow us to determine the cut-off points for fee type classification. To determine these fee type cut-off points, we use historical data on university tuition fees by Canadian provinces from Statistics Canada [1]. These historical

[1] We combine the data series 3710004501 and 3710016001 to get full coverage between 1985 and 2019.

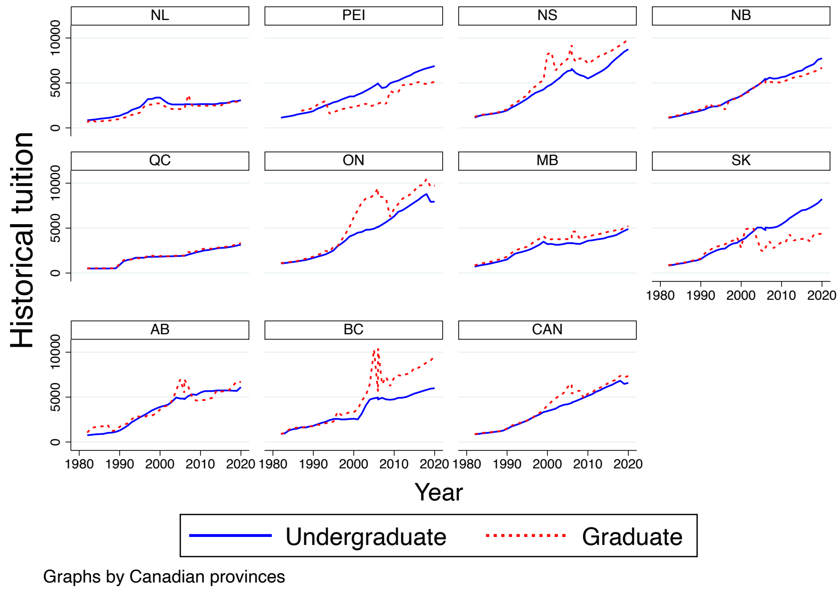


Figure 1: Historical university tuitions by Canadian provinces

records report the average amount of tuition paid by a full-time, Canadian ("domestic") student enrolled in university for each Canadian provinces, at both the undergraduate and the graduate level[2]. Figure 1 plots these tuition fees by Canadian provinces. These series allows us to determine the lower and upper cut-off values used to distinguish between college, university and other PSEs. Moreover, differences between undergraduate and graduate tuition allow us to separate students pursuing a Bachelor's degree from those enrolled in post-graduate programs.

The classification of fee types relies on two kinds of dispersion observed in the tax variable `TUTDN`. The first is the dispersion of fees *across* types of PSE institution: colleges generally charge lower tuitions than universities, while private PSE institutions would charge significantly more than publicly-funded ones. Finnie and Pavlic (2013) exploit these two peaks in tuition paid (one for college and one for university) to determine the fee types for each individual, a method we follow here. The second type is the dispersion *within* university tuitions, namely between undergraduate and graduate levels. Although the fee gaps are small between them, it is still possible to distinguish between these two levels when combined with PSE attendance histories . For example, suppose we observe a group of individuals paying university-level tuition for six consecutive years. If they were

[2]It is possible to further distinguish between fields of study using a different historical tuition series from Statistics Canada, although there are no existing data providing the full details on tuition fees.

Table 1: Types of inferred Canadian post-secondary education: full time students

PSE classification	Consecutive attendance	Fee type	Previous PSE
No Canadian PSE	-	-	-
College "drop-out"	$<$ 22 months	$<$ undergraduate fee	-
College degree	$\geq$ 22 months	$<$ undergraduate fee	-
University "drop-out"	$<$ 30 months	undergraduate fee	-
University degree	$\leq$ 30 months	undergraduate fee	-
Post-graduate	$\geq$ 12 months	graduate fee	university degree

enrolled as full-time student for eight months per year in the first four years, followed by 12-month enrolments in the last two years, this pattern is consistent with a typical "Bachelor's-Master's" progression without a gap year in between. The presence of gap years does not complicate the classification process substantially, since the sequence of PSE attendance events provides most of the relevant information[3]. Using joint information from attendance records and the fee types, we can classify individuals into one of five PSE types as illustrated in Table 1. These proxies for Canadian PSE attainment are used to capture the deterministic part of labour income and the initial heterogeneity in persistent income component when they first appear in our sample.

We apply a similar procedure for individuals who report tuition fees and education deductions as part-time students. No special adjustments are needed to recover the number of months enrolled, since part-time students claim deductions in the same way as full-time students, with differences only in the per month amounts. Historical tuition data also include average part-time tuition fees by Canadian provinces, which we use to identify the fee categories following the same logic as above. Classification cut-off values are adjusted to account for the longer time typically required for part-time students to complete their programs. Finally, individuals who never claimed any PSE related deductions or tax credits are classified as having "no Canadian PSE", since we cannot distinguish between those with no PSE experiences and those who completed their entire PSE outside Canada.

[3]This classification approach may underestimate the highest level of education attainment if an individual only attended graduate programs in Canada. For example, if an individual obtained an undergraduate degree from overseas, then attended a PhD program in Canada, then she would only leave a five-year PSE attendance record with a university fee type. Depends on how similar the PhD and undergraduate tuitions are, our approach may not able to tell apart whether she attended a PhD program or an undergraduate program which took five-years to complete.

3.3 Sample creation

We construct three types of samples to address distinct research questions regarding the evolution of income dynamics among Canadian-born working households. The first set of samples, which we refer to as the *historical samples*, is designed to examine whether the nature of income dynamics has changed since the 1980s despite substantial variation in aggregate economic conditions and government policies. To do so, we exploit the long data availability of LAD and consider a sampling period of 1985-2019, spanning 35 years. We partition this period into five equal-length sub-periods: 1985-1991, 1992-1998, 1999-2005, 2006-2012 and 2013-2019, each containing seven years ($T = 7$). The choice of this sub-period length reflects a tradeoff between generating a sufficient number of distinct periods for comparison and maintaining a good numerical performance of the estimation algorithm, as existing evidence suggest that both very short and very long panels can lead to performance deterioration. Since our estimation requires a statistical sampling method applying at the individual-year level, we limit the sample size to 5000 randomly-selected observations ($N = 5000$) to keep computational time manageable. This procedure yields six balanced panels covering 35-year period. We do not impose restrictions on demographic variables (such as age or gender) or geographic location, as our goal is to capture the time evolution of the Canadian working population as accurately as possible. Descriptive statistics on the pre-tax labour incomes over the period of 1985-2015 are provided in Bowlus et al. (2022).

The second type of sample, which we term the *life-cycle* sample, is designed to study how income dynamics evolve over a working family's life-cycle. To this end, we focus on the 1964 birth cohort and follow its members over their entire working lives from 1989 to 2018, corresponding to ages 25-54. We divide this 30-year sampling window into five equal-length sub-panels: 25-30 (1989-1994), 31-36 (1995-2000), 37-42 (2001-2006), 43-48 (2007-2012) and 49-54 (2013-2018). The first two sub-panels capture the early stages of working life, while the final sub-panel cover the prime-aged years. Our focus on the 1964 cohort is motivated by the reliability of imputing post-secondary education (PSE) from tax records. Individuals born in 1964 turn 18 in 1982, the first year for which LAD data are available, allowing us to observe their full PSE window (beginning at age 18) and impute educational attainments without data loss due to un-observability. Earlier cohorts born prior to 1964 would have incomplete deduction and tuition tax credit histories, rendering

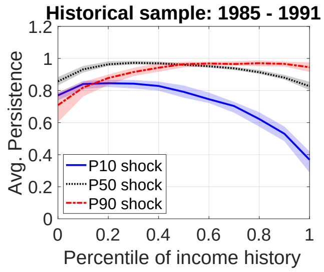
PSE imputation less reliable. We limit the sample size to 5000 by randomly selecting individuals who satisfy the following criteria: they were born in 1964, earned positive labour income, filed tax returns in every year from 1989 to 2018, and are continuously present throughout the 30-year window. No additional demographic or geographic restrictions are imposed on this type of sample. We present descriptive statistics for the life-cycle sample in Table 2 of the appendix.

To assess the representativeness of the 1964 cohort, we complement the life-cycle sample with a third type of sample, referred to as the *cohort sample*. These samples are constructed in similar ways as the life-cycle sample but for the following additional cohorts: 1971, 1978 and 1985. These birth cohorts are spaced seven years apart to ensure that the calendar time corresponding to a given age range does not overlap across cohorts. For instance, the ages 25-31 sub-panel for the 1971 birth cohort spans 1996-2002, which does not overlap with the corresponding period for the original 1964 cohort (which is 1989-1994 in calendar time). This non-overlapping sub-panel design ensures that same aggregate or time-dependent economic conditions do not affect two birth cohorts simultaneously at similar ages. This sampling strategy produces the following age panels: "late 20s to early 30s" (available for birth year 1964, 1971, 1978 and 1985), "mid 30s to late 30s" (available for birth year 1964, 1971 and 1978), "late 30s to mid 40s" (available for birth year 1964 and 1971 only). We set the sample size to $N = 5000$ as for the life-cycle panels while lifting the continuous presence requirement for each of the age-cohort sub-panels.

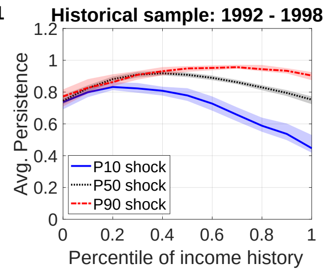
Across all three types of samples, we impose a common set of eligibility criteria. Individual workers leading the household must born in Canada and have no missing or imputed information on pre-tax labour income or basic individual and family demographics. Furthermore, individuals included in the historical and cohort samples must not be deceased during the relevant sampling periods, while those in the life-cycle sample must remain alive throughout their working years.

4 Results

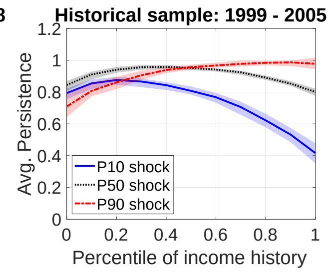
We estimate the nonlinear income models described by Eq 10 and Eq 11 separately on each of the sample types discussed in the previous section. These estimates capture nonlinear income persistences in form of derivative (marginal) effects. We begin by presenting results based on historical samples, first examining each panel separately to highlight time-specific features, then



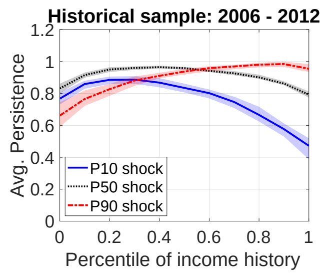
(a) 1985 - 1991



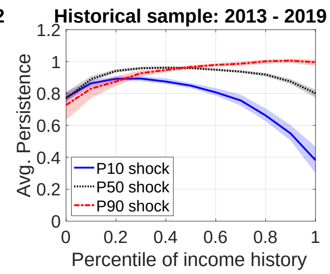
(b) 1992 - 1998



(c) 1999 - 2005



(d) 2006 - 2012



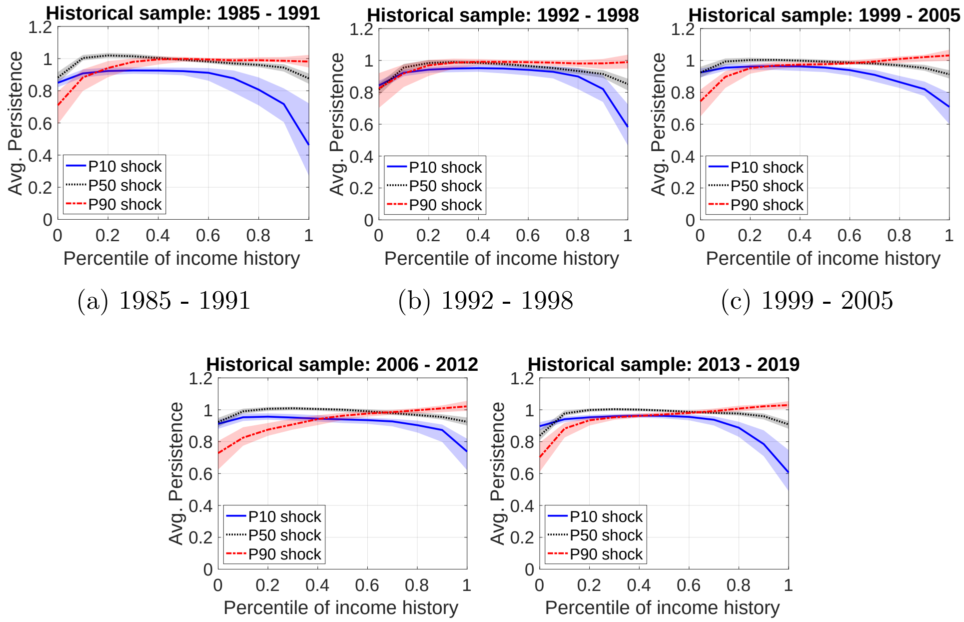
(e) 2013 - 2019

Figure 2: "With-in" results: average persistence of y_t by shock percentiles

comparing results across panels to assess the temporal stability of income dynamics. We next turn to the life-cycle sample to investigate how income persistence evolves as workers age. Finally, we compare results across birth cohorts to assess the robustness of Canadian income dynamics over time and remark on the recessionary effects.

4.1 Historical sample

We organize the discussion of results using a "within-between" structure. First, we present the within-period estimates based on each sub-period of the historical sample, following by a cross-period comparison. Results based on residual log income are presented in Figure 2. Each panel corresponds to one of the historical sub-periods in our sampling window, and illustrates two key dimensions of heterogeneity in income persistence. The x-axis plots the percentile of income history z_{t-1}, where lower (higher) percentile is associated with a lower (higher) income levels in the previous period. The y-axis reports the estimated persistence of residual log income. Three curves in each panel represent different percentiles of current income shock: a P_{10} shock correspond to a large negative ("bad") shock, a P_{50} shock represents a small income shock with value closer to zero, and a P_{90} shock represents a large positive ("good") shock.

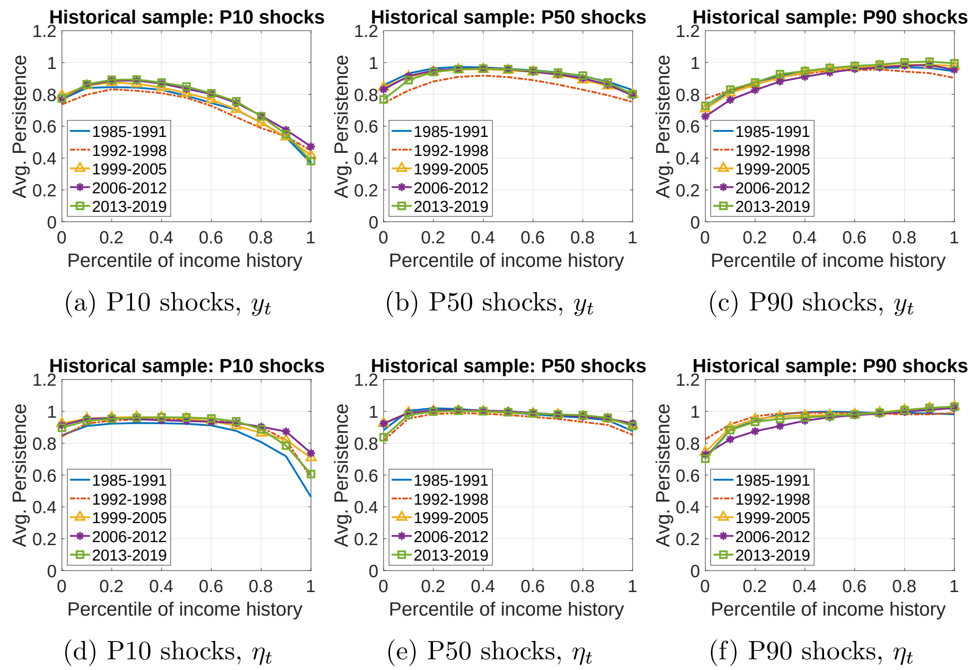


(a) 1985 - 1991 (b) 1992 - 1998 (c) 1999 - 2005

(d) 2006 - 2012 (e) 2013 - 2019

Figure 3: "With-in" results: average persistence of η_t by shock percentiles

Within each historical period, we found that shocks of the same rank (percentile) can exhibit very different degrees of persistence depending on the individual's income history. For households with high income households (at or above P_{50} of income distribution), large positive income shocks are highly persistent, while the adverse effect from large negative shocks dissipate rapidly. The persistence associated with median shocks falls between these two extremes. In other words, the persistence of residual log income is monotonically increasing in the rank of current income shocks, but only among households in the upper part of the income distribution. The 95% non-parametric bootstrap confidence intervals for persistence estimates associated with different shocks do not overlapping in any of the five historical sub-periods, leading strong statistical support to this finding. For households with very low income histories (at or below P_{10} of income distribution), differences in income persistence across shock magnitudes are much less pronounced. The 95% bootstrap confidence intervals overlap substantially in nearly all historical sub-periods, indicating that persistence is more homogenous in this part of the income distribution. This result suggests that canonical income process correctly capture the relative homogeneity of persistence among low-income households, even if it captures the incorrect magnitude. For households with income histories between the first and second quartiles, income persistence follow a inverted-U shape,



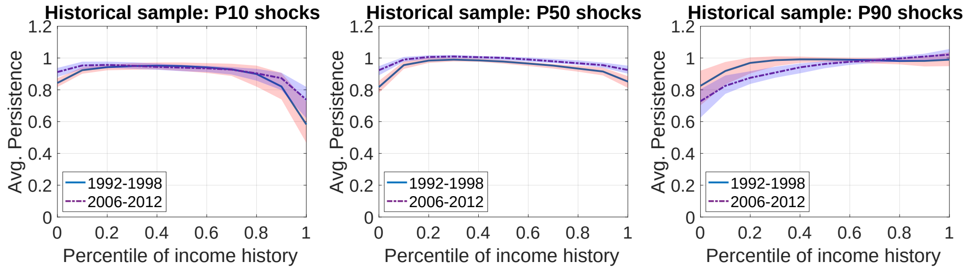
(a) P10 shocks, y_t (b) P50 shocks, y_t (c) P90 shocks, y_t

(d) P10 shocks, η_t (e) P50 shocks, η_t (f) P90 shocks, η_t

Figure 4: "Between" results: average persistence by shock percentiles

with shocks of intermediate magnitude (closer to the median of shock distribution) having the most persistent effects. Taken together, these results indicate that only households with residual log income above the third quartile experience the most favourable conditions for income growth, driven by the strong persistence of positive shocks.

The relatively lower persistence in residual log income observed earlier is partly attributable to the influence of transitory income shocks occurring in each period. To isolate the impact of these transitory components, we turn to estimates based solely on the persistent income components. See Figure 3. Compared to results based on residual log income, several patterns documented previously continue to hold. In particular, the monotonic relationship between income persistence and shock percentiles among high income households remains evident, although the differences across shock types are somewhat attenuated. For low income households, however, the earlier finding of homogeneous persistence is no longer valid (with the exception of the 1992-1998 period). In this part of the income distribution, P_{90} shocks are now found to be less persistent than lower ranked shocks. Moreover, for households with persistent income histories between the 20th and 70th percentile, variations in income persistence across shock magnitudes become much less pronounced with estimated persistence consistently above 0.9 and often close to unity. This suggests that



(a) P10 shocks (b) P50 shocks (c) P90 shocks

Figure 5: Average persistence of η_t with recessionary periods

canonical income process assuming homogenous income persistence may adequately describe income dynamics for these middle-income households, but they fail to capture the heterogeneity present at the lower and upper ends of the persistent income distribution. A striking difference between Canadian income dynamics and those documented for the United States concerns the persistence of adverse income shocks among low income households. In our estimates, the persistence associated with negative shocks in this group is around the low to mid-0.9 range, with confidence intervals that do not include unity. This implies that although adverse shocks are relatively persistent, they still decay over time, allowing upward income mobility to gradually increase. This stands in contrast to the unit-root behaviour of large negative shocks often reported for low income households in the U.S., which has been cited as a potential driver of “low income traps”, such that they remain persistently disadvantaged over their lifetimes.

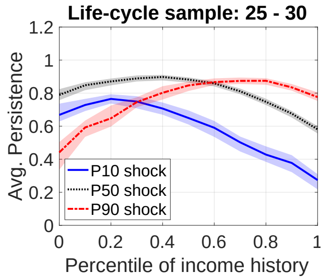
It is also instructive to examine the varying width of the bootstrap confidence intervals across shock percentiles. These bands provide two key insights. First, our heterogenous income persistence estimates are very precise for households experiencing shocks near the median of the distribution. Second, the widening of the confidence interval at the tails (i.e., for P_{10} or P_{90} shocks) indicates that the exposure to more extreme shocks is not uniform across individuals. In particular, the likelihood of receiving a large positive shocks appears to increase with the rank of current persistent income.

Comparing results across all five historical sub-periods reveals that the features of income dynamics discussed above are remarkably stable over time and can therefore be regarded as robust characteristics of Canadian labour income. Figure 4 contrasts estimates based on residual log incomes (top panels) with those based on persistent incomes (bottom panels). Across all sub-periods, income persistence patterns display a high degree of consistency. The persistence of median shocks shows the least variation across periods (with the exception of 1992-1998), whereas more extreme shocks exhibit slightly larger period-to-period fluctuations. Notably, the persistence of very low rank income shocks (i.e. very adverse events) remains strictly below unity for all income histories, underscoring the limited long-term impact of such bad events. A comparison between the two panels also highlight the extent and distribution of transitory shocks: these shocks appear predominantly affect households in the neighbourhood (immediately above and below) of median income history, depending on the type of current shocks.

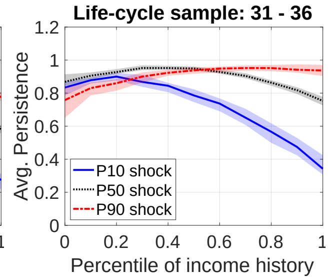
Our sample period includes two major recessionary episodes in Canada: the early 1990s recession and the Great Recession of 2008. Panel (d) of Figure 4 indicates a mild increase in the persistence of negative shocks among high income households during the sampling period overlapping with the Great Recession. Panel (e) shows a decline in decline in the persistence of median shocks during the early 1990s recession. Figure 5 focus on the sub-periods encompassing these two recessions and provide additional details on their differential impacts. The sub-populations of individuals most affected by these recessionary events differs across episodes: the early 1990s recession is associated with a decrease in the persistence of median shocks (panel (b)), whereas the Great Recession increased the persistence of negative income shocks for both low and high income households (panel (a)), and reduced the persistence of positive shocks for households with below-median income histories (panel (c)).

4.2 Life-cycle sample

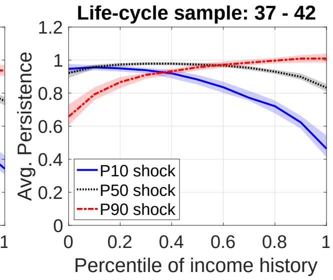
Results from the historical sample provide an aggregate overview of income dynamics among Canadian working households across different time periods. We now turn to a more granular perspective by examining how income shock persistence evolves over a worker's life-cycle. This analysis addresses a key question: how does the income process change as workers age? Figure 6 presents results based on residual log income. Panel (a) shows that income persistence among young workers aged between 25 to 30 is substantially lower than the "population average" (i.e. the estimates obtained when pooling individuals of all ages) for the 1964 birth cohort. Notably, the persistence of negative income shocks is particularly low during this early career stage, valued around 0.7, but grows as workers age. This indicates that while low-percentile income shocks are more likely to occur at the beginning of working life (possibly due to job changes), their limited persistence prevents



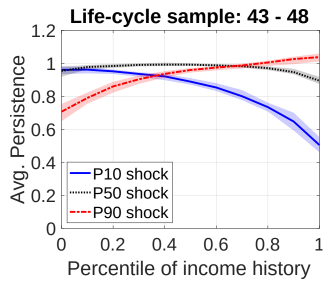
(a) Age 25 - 30



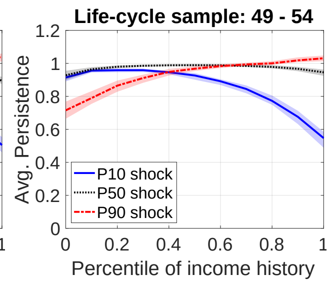
(b) Age 31 - 36



(c) Age 37 - 42



(d) Age 43 - 48

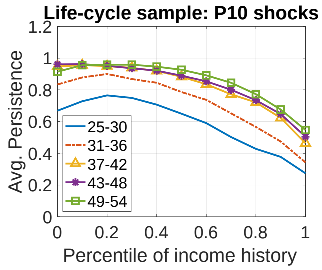


(e) Age 49 - 54

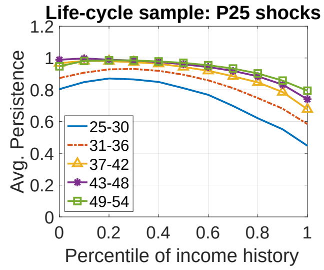
Figure 6: "With-in" results: average persistence of y_t by shock percentiles

them from exerting lasting effects on labour income. Equally striking is the very low persistence of favourable income shocks for young workers (below 0.5). Although such low persistence implies that positive shocks contribute little to long-run income growth, the dynamic effect of persistence suggests that prior adverse income events dissipate more quickly following the arrival of positive shocks. This feature is crucial for younger workers: a positive income shock may not generate substantial cumulative income gains, but it helps "wipe out" the lingering effects of earlier negative shocks. The reverse dynamic holds for high-income young workers as well: a negative shock may have only short-term effects on income but may nonetheless erode the upward trajectory created by previous favourable shocks.

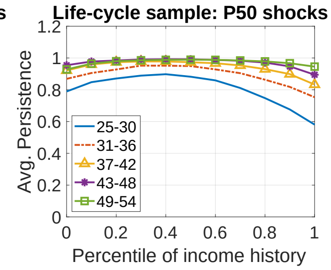
As workers enter their 30s (Panel (b)), average income persistence increases further, particularly for higher-ranked shocks. Low income households exhibit a period of relatively homogeneous income persistence, suggesting that the dynamic effects of shocks become less pronounced than during their late twenties. This may partially due to them start climbing the job ladder within their current jobs so that any income shocks arrived would likely to stay around longer, as workers have lower incentives to switch jobs years into their career. Households with more favourable income histories experience only minor changes in income dynamics during this stage. As individuals progress into



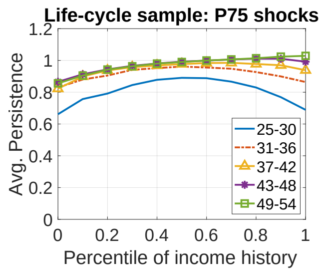
(a) P10 shocks



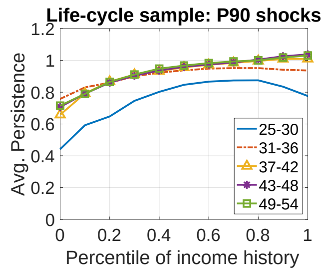
(b) P25 shocks



(c) P50 shocks



(d) P75 shocks

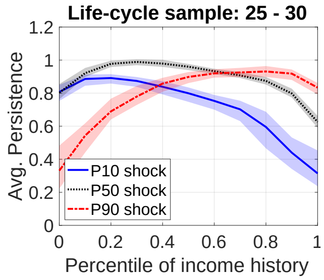


(e) P90 shocks

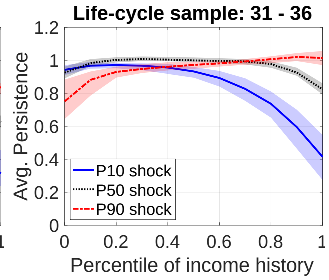
Figure 7: "Between" results: average persistence of y_t by shock percentiles

their late 30s, 40s and early 50s, changes in the income dynamics becomes more subtle. Income persistence begins to converge for households with income history around the 40th percentile. For this income group, the gap between the persistence of median and more extreme income shocks gets narrower over time as income dynamics start to stabilize. Figure 7 provides a direct comparison across age groups. Panel (a) to (c) reveal a clear increase in the shock persistence as workers age, but this growth is concentrated among shocks with rank below or at the median. There is limited evidence of increasing persistence for higher ranked income shocks, and shock persistence becomes largely independent of individual worker's age once they reach their mid-40s.

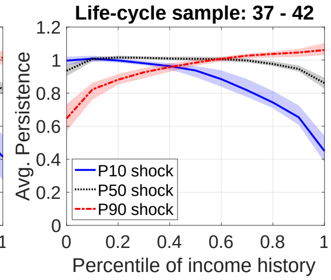
We next turn to results estimated using only the persistent income components. See Figure 8. Comparing results between residual log income and persistent income highlights three important differences: first, the majority of life-cycle patterns identified from the residual log income still holds when we look at the core component of residual income, after removing shocks with contemporaneous effects only, with the exception of the homogenous shock persistence observed among low income household aged 31-36. Second, the impact of transitory shocks on income process diminishes as workers age, which are indicated by the shrinking gaps between the two sets of results. This suggests that exposure to different types of income risk evolves over a worker's



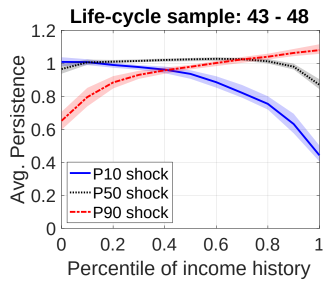
(a) Age 25 - 30



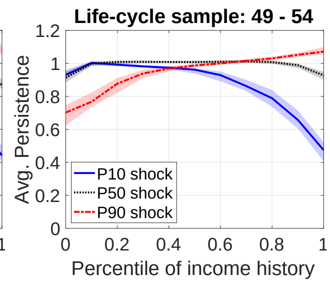
(b) Age 31 - 36



(c) Age 37 - 42



(d) Age 43 - 48

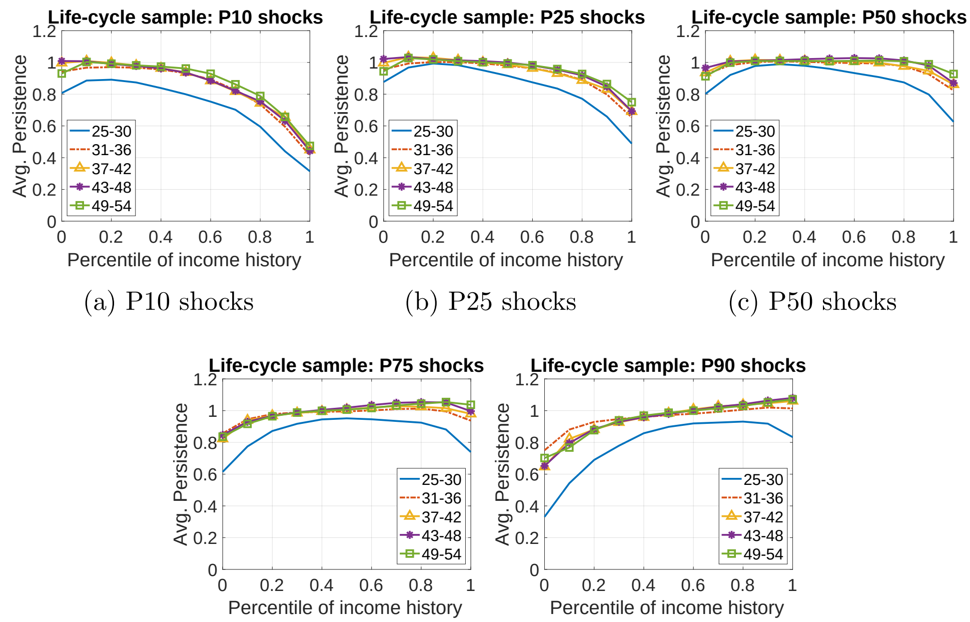


(e) Age 49 - 54

Figure 8: "With-in" results: average persistence of η_t by shock percentiles

life-cycle. Third, we find that low income households from the 1964 birth cohort experienced an extended period of "permanent" adverse income shocks during their working years. Panel (c) and (d) show point estimates of income persistence very close to unity during the age period of 37 to 48. This finding has significant implications for worker welfares: low income households face heightened risks of prolonged periods of reduced income levels and limited upward mobility, as the permanence of adverse shocks exerts sustained downward pressure on income growth. Lastly, 9 illustrates the evolution of income persistence by income history and income shock types over the life-cycle. A mild "fanning out" patterns on shock persistences is apparent: households with income histories near the median of income distribution would experience a relatively stable income process from their 30s and onward, whereas households with more extreme income histories would experiences more noticeable variations in persistence as they age.

Overall, our life-cycle analysis underscores the limitations of canonical income process. We document strong history and state dependence in income persistence, and the form and extent of these nonlinear dependencies evolve as workers age. Consequently, assuming a fixed income persistence parameter risks overstating or understating the cumulative effects of income shocks, as such assumptions ignore the dynamic effects of income shocks.



(a) P10 shocks (b) P25 shocks (c) P50 shocks

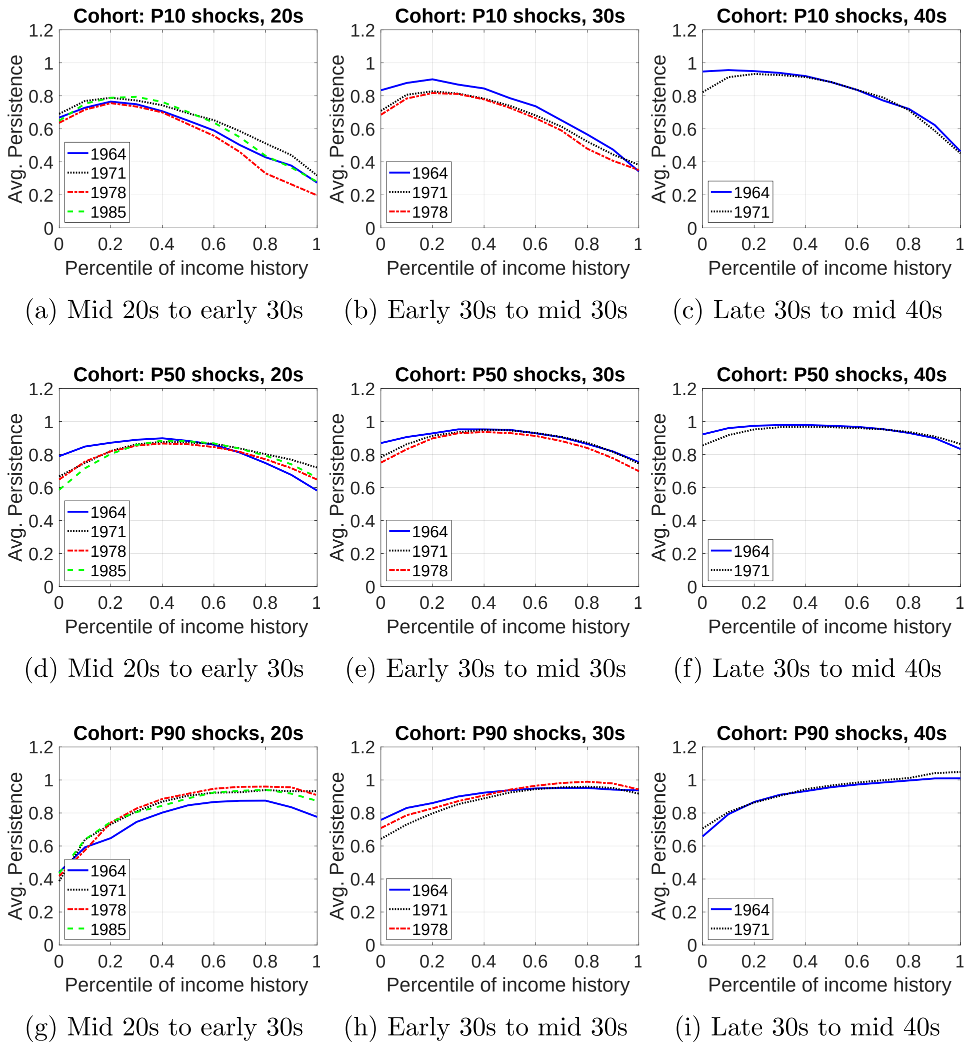
(d) P75 shocks (e) P90 shocks

Figure 9: "Between" results: average persistence of η_t by shock percentiles

4.3 Cohort sample

An important question is whether the income dynamics documented previously are unique to the 1964 birth cohort or generalize well to other cohorts of Canadian-born workers. To assess the representativeness of our findings, we re-estimate the nonlinear income models using samples of Canadian working households selected under similar criterions as the life-cycle sample, but drawn from different birth cohorts. We organize the results by age and categorize workers into three groups: those in their 20s (mid to late 20s), 30s (early to late 30s), and 40s (up to the mid-40s).

Figure 10 presents the results based on residual log incomes. Each row corresponds to a particular type of shocks (in ascending order of percentiles) and each column corresponds to a particular age group. For workers in their 20s, we observe more substantial between-cohort variations in income persistence, with the extent of dispersion depending on the rank of the shock. Notably, for favourable income shocks, the 1964 cohort exhibits lower persistence than later cohorts, suggesting that upward income mobility may have been weaker for this cohort relative to those born seven to twenty-one years later. By contrast, the average persistence of median shocks during the 20s remain statistically indistinguishable across cohorts. There is no systematic pattern in the dispersion of negative shock persistence across these four birth cohorts. Overall, the difference in income



(a) Mid 20s to early 30s (b) Early 30s to mid 30s (c) Late 30s to mid 40s

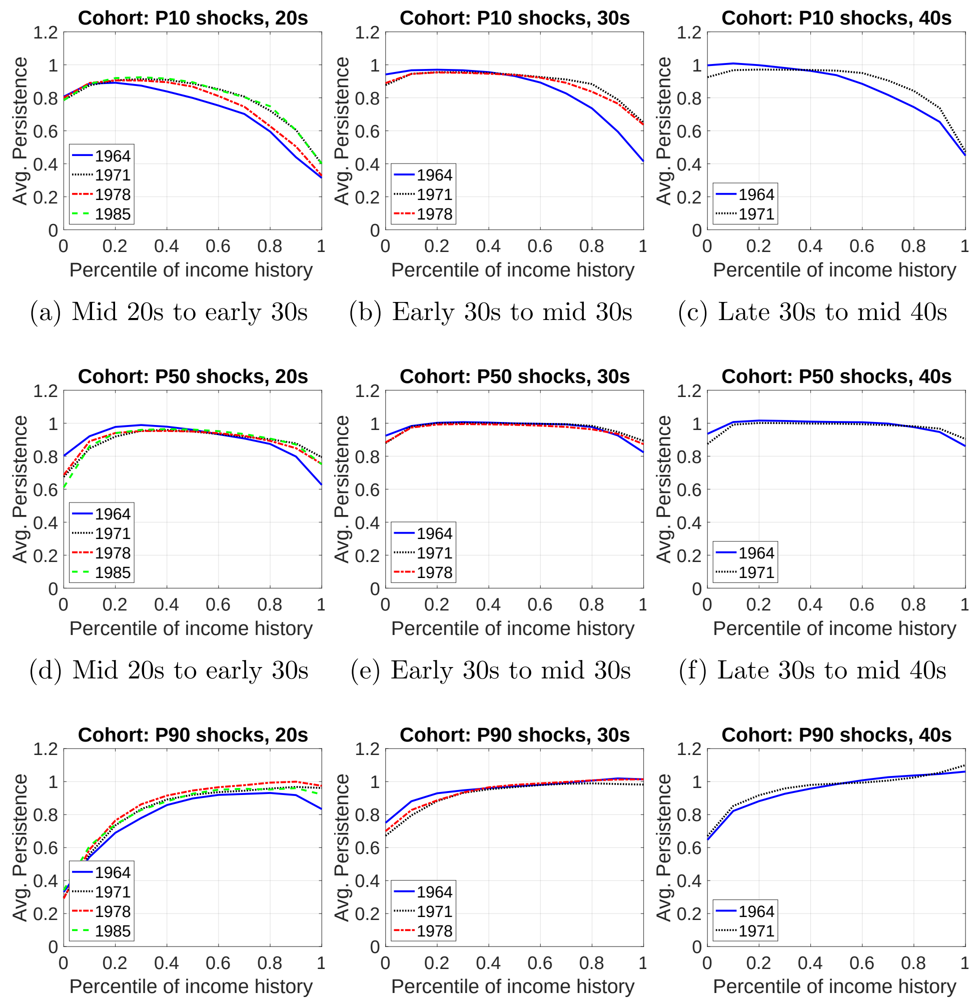
(d) Mid 20s to early 30s (e) Early 30s to mid 30s (f) Late 30s to mid 40s

(g) Mid 20s to early 30s (h) Early 30s to mid 30s (i) Late 30s to mid 40s

Figure 10: Average persistence of y_t by birth cohort

dynamics for young workers in their 20s are broadly similar across cohorts, with the main exception occurring in response to large positive shocks. Moving into their 30s, the 1964 cohort experienced a stronger persistence for low percentile income shocks relative to later cohorts, while differences for other types of shocks remain marginal. These between-cohort differences further diminishes in the 40s, where patterns of income persistence remains very similar between 1964 and 1971 birth cohorts.

Our results also shed light on how exposure to macroeconomic conditions, particularly recessions, shapes income dynamics. Several cohorts in our sample experienced significant recessions during their working lives: those born in 1971 and 1985 encountered recessions in their 20s, those born in 1971 and 1978 faced recessions in their 30s, and only the 1964 cohort experienced a reces-



(a) Mid 20s to early 30s (b) Early 30s to mid 30s (c) Late 30s to mid 40s

(d) Mid 20s to early 30s (e) Early 30s to mid 30s (f) Late 30s to mid 40s

(g) Mid 20s to early 30s (h) Early 30s to mid 30s (i) Late 30s to mid 40s

Figure 11: Average persistence of η_t by birth cohort

sion in their 40s. These recessionary effects are especially evident in the results based on persistent income components (Figure 11). For workers in their 20s, the 1971 and 1985 birth cohorts experienced a higher persistence for negative shocks relative to others (panel (a)). Similarly, workers in their 30s born in 1971 and 1978 experienced moderate yet asymmetric impacts depending on their income history: low income households faced a lower persistence for large positive shocks (panel (h)), while high income households experienced a stronger persistence for large negative shocks (panel (b)). The 1964 cohort, who were in their mid-40s during the Great Recession, displayed a distinct pattern: the income process for low income households followed a unit root process (panel (c)), while households with higher incomes appeared to be relatively less affected during this period. Due to data limitations, we are unable to compare cohorts beyond their late 40s. Finally, it is worth noting that these recessionary effects are less pronounced in results estimated using residual log incomes, largely because transitory income shocks introduced additional noise during the recessions.

5 Conclusion

We study how income dynamics of Canadian-born workers evolves over time and across their life-cycle. Our analysis focuses on pre-tax household ("parental") labour income, from which we construct two different income measures for estimating income persistence, namely the residual income in natural logarithms and the persistent income net of transitory income fluctuations. A novel feature of our empirical approach is the construction of a proxy for Canadian post-secondary education attainment using information on education deductions and tuition tax credits from individual tax records. This proxy allows us to control for the influence of higher education on the deterministic component of labour income, an important dimension that previous studies have been unable to capture.

Using a flexible nonlinear income model and longitudinal administrative income data derived from tax records, we find that the persistence of household labour income exhibits strong state and history dependence. Results from the historical samples indicate that favourable income shocks are long-lasting only for households with high income histories, whereas adverse shocks have no permanent effects regardless of household income histories. This contrasts with evidence from the United States, where low rank income shocks are often found to be permanent. Our findings suggest that low income working households in Canada likely experienced greater upward income mobility during the 1985-2019 period, compared to their American counterparts. Moreover, the non-permanence of adverse income shocks is a robust feature across both income measures we consider and remains stable over the three decades of our sample.

Turning to the life-cycle dimension, evidence from a panel of workers born in 1964 shows that income persistence increases substantially during the first decade of the working life. Persistence is relatively low between ages 25 and 34, implying two features: (i) adverse income events occurred early in life have limited long-term impacts, and (ii) the dynamic persistence effects of favourable income shocks accelerates the decay of pre-existing negative income shocks. As workers enter their late 30s, their income processes stabilize with a higher average persistence than their first working decade. In later working years, we observe a narrowing gap in income persistence across various ranks of shocks, particularly for households around the median of income distribution. Evidence from subsequent birth cohorts confirms that these life-cycle patterns are consistent across generations of Canadian-born workers. We also uncover clear signs of recessionary effects on income dynamics. These include weaker persistence of favourable shocks among low-income households and stronger persistence of adverse shocks among high-income households during recessionary episodes. Taken together, our findings emphases the importance of employing a flexible income process: many of the subtle features documented here, such state and history dependence, dynamic persistence effects and cohort-specific patterns, are abstracted away in canonical income process specifications.

References

Arellano, M., Blundell, R., & Bonhomme, S. (2017). Earnings and consumption dynamics: a nonlinear panel data framework. *Econometrica*, *85*(3), 693–734.

Arellano, M., Blundell, R., Bonhomme, S., & Light, J. (2024). Heterogeneity of consumption responses to income shocks in the presence of nonlinear persistence. *Journal of Econometrics*, *240*(2), 105449.

Baker, M., & Solon, G. (2003). Earnings dynamics and inequality among canadian men, 1976–1992: Evidence from longitudinal income tax records. *Journal of Labor Economics*, *21*(2), 289–321.

Blackburn, M. L., & Bloom, D. E. (1993). The distribution of family income: measuring and explaining changes in the 1980s for canada and the united states. In *Small differences that matter: Labor markets and income maintenance in canada and the united states* (pp. 233–266). University of Chicago Press.

Bonikowska, A., et al. (2024). *Cumulative earnings of black, chinese, south asian and white individuals born in canada* (Tech. Rep.). Statistics Canada, Analytical Studies and Modelling Branch.

Bowlus, A., Gouin-Bonenfant, É., Liu, H., Lochner, L., & Park, Y. (2022). Four decades of canadian earnings inequality and dynamics across workers and firms. *Quantitative Economics*, *13*(4), 1447–1491.

Finnie, R., & Pavlic, D. (2013). Background characteristics and patterns of access to postsecondary education in ontario: Evidence from longitudinal tax data.

Fortin, N. M., & Lemieux, T. (2015). Changes in wage inequality in canada: An interprovincial perspective. *Canadian Journal of Economics/Revue canadienne d'économique*, *48*(2), 682–713.

Gottschalk, P. (1993). Changes in inequality of family income in seven industrialized countries. *The American Economic Review*, *83*(2), 136–142.

Kitao, S., Suzuki, M., & Yamada, T. (2025). *Nonlinear earnings dynamics and inequality over the life cycle: Evidence from japanese municipal tax records* (Unpublished doctoral dissertation). Tohoku University.

Morissette, R., & Berube, C. (1996). Longitudinal aspects of earnings inequality in canada. *Statistics Canada WP*, *94*.

Pendakur, K., & Pendakur, R. (2007). *Colour my world: Have earnings gaps for canadianborn ethnic minorities changed over time?* Toronto: University of Toronto Press.

Picot, G. (2001). Working time, wages, and earnings inequality among men and women in canada, 1981-1993. *Working Time in Comparative Perspective*, *1*, 109–143.

Richardson, D. H. (1997). Changes in the distribution of wages in canada, 1981-1992. *Canadian Journal of Economics*, 622–643.

Zhang, X. (2014). *What can we learn about low-income dynamics in canada from the longitudinal administrative databank?* Statistics Canada/Statistique Canada.

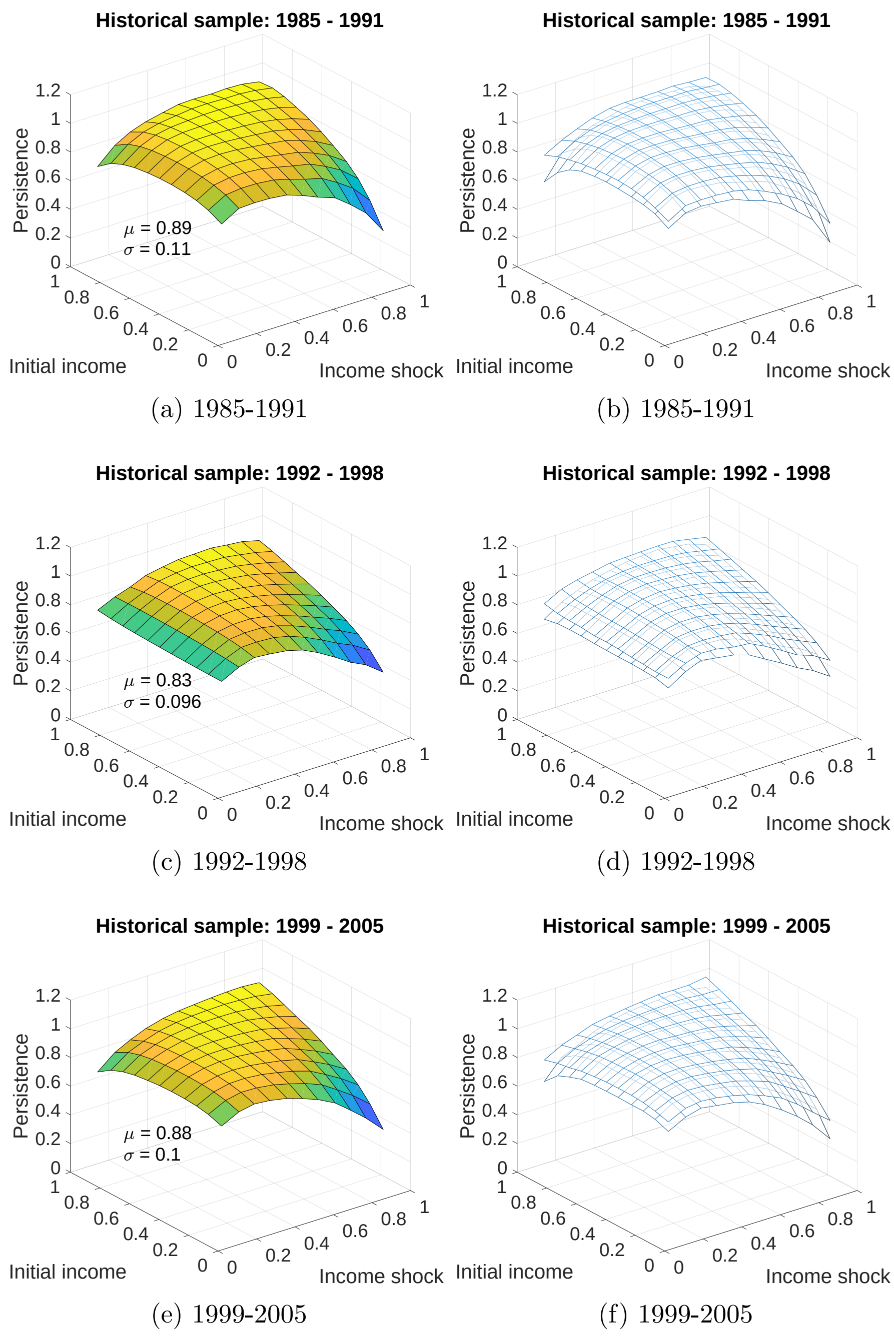
A Descriptive statistics: life-cycle sample

Table 2: Descriptive statistics of labour income: life-cycle sample

Time periods	1989-1994	1995-2000	2001-2006	2007-2012	2013-2018
Age groups	25-30	31-36	37-42	43-48	49-54
P_5	45400	55100	63700	70200	71500
P_{10}	51500	63400	74200	82100	84100
P_{25}	65000	80500	95100	106400	109300
P_{50}	85100	103500	123200	140500	146500
P_{75}	108600	131200	157700	182300	193100
P_{90}	131400	160300	194900	228700	245300
P_{95}	145800	180400	222000	262500	285500

B Additional results

B.1 Historical sample: residual log labour income



(a) 1985-1991

(b) 1985-1991

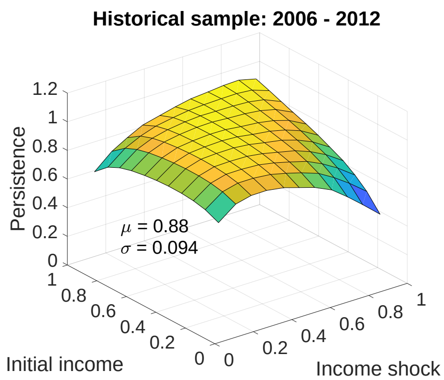
(c) 1992-1998

(d) 1992-1998

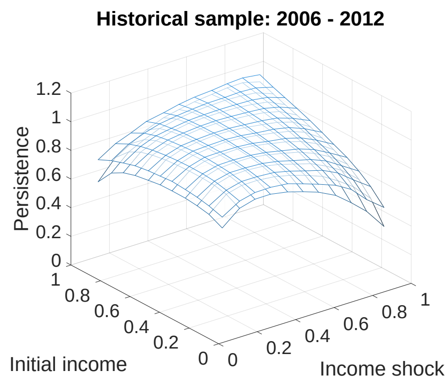
(e) 1999-2005

(f) 1999-2005

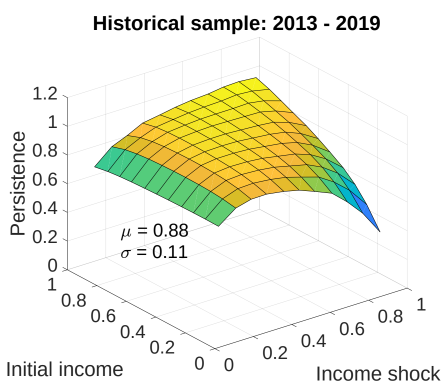
Figure B.1: Income persistence of y_t by historical periods



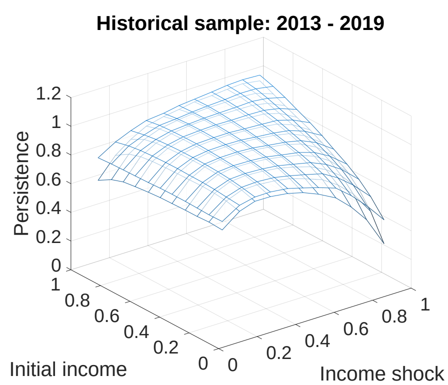
(a) 2006-2012



(b) 2006-2012



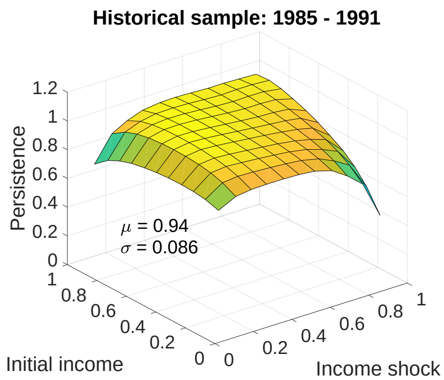
(c) 2006-2012



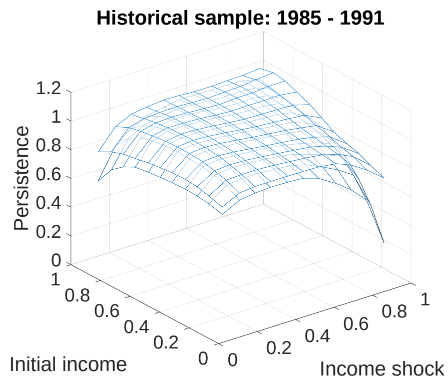
(d) 2006-2012

Figure B.2: Income persistence of y_t by historical periods

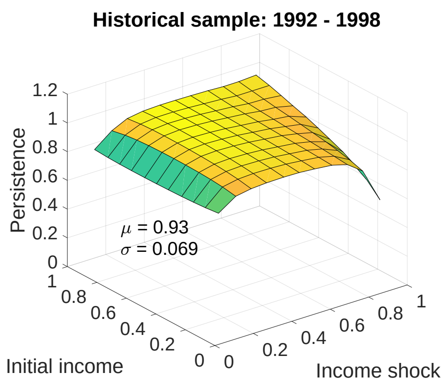
B.2 Historical sample: persistent income



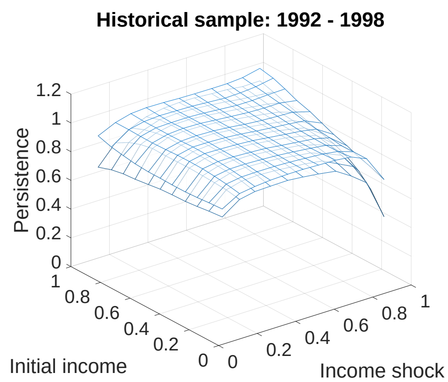
(a) 1985-1991



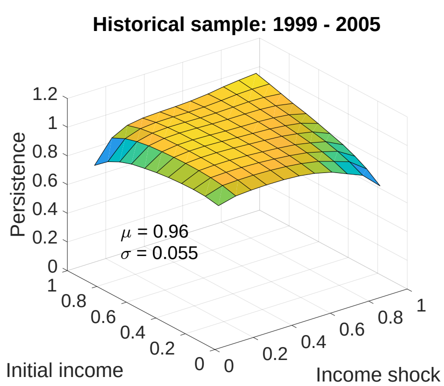
(b) 1985-1991



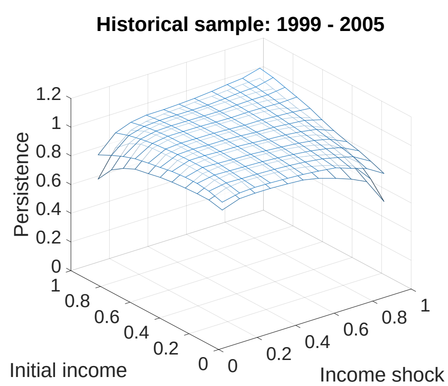
(c) 1992-1998



(d) 1992-1998

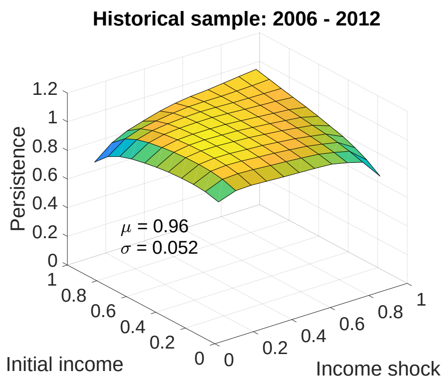


(e) 1999-2005

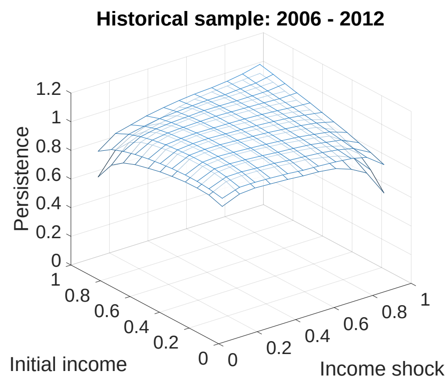


(f) 1999-2005

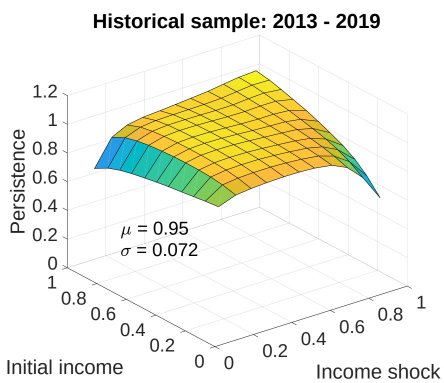
Figure B.3: Income persistence of η_t by historical periods



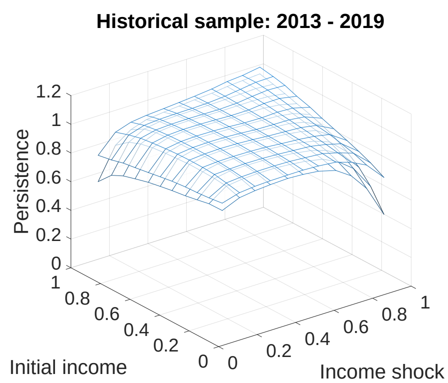
(a) 2006-2012



(b) 2006-2012



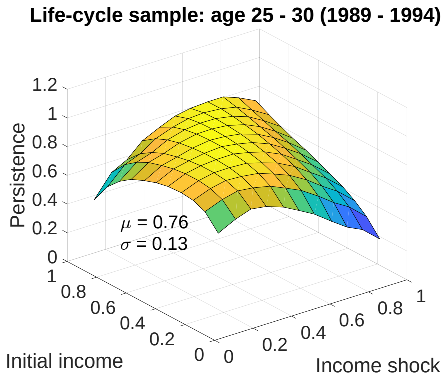
(c) 2013-2019



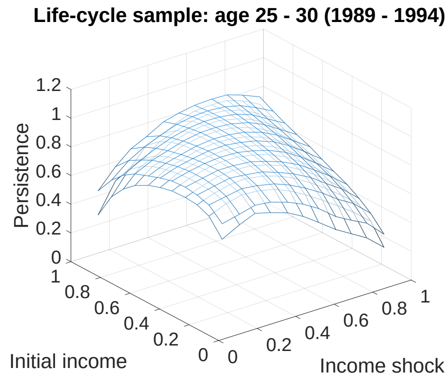
(d) 2013-2019

Figure B.4: Income persistence of η_t by historical periods

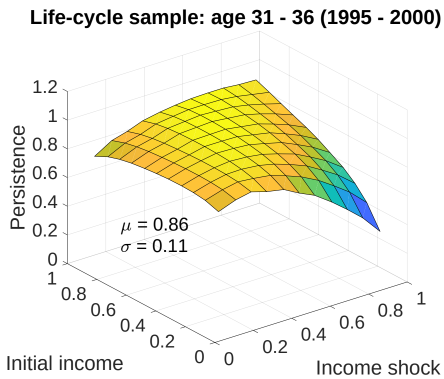
B.3 Life-cycle sample: residual log labour income



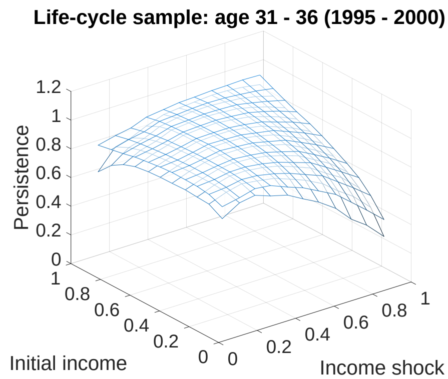
(a) 1964 cohort: 25-30



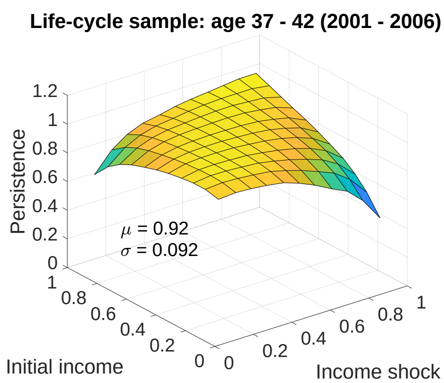
(b) 1964 cohort: 25-30



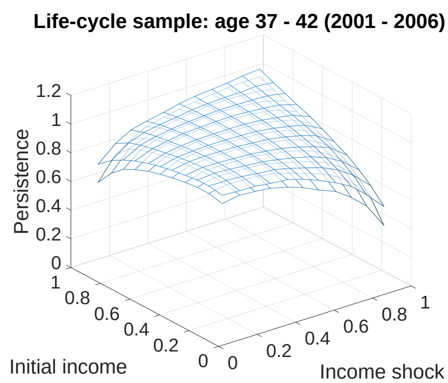
(c) 1964 cohort: 31-36



(d) 1964 cohort: 31-36

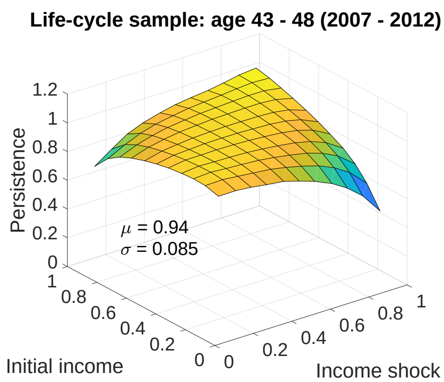


(e) 1964 cohort: 37-42

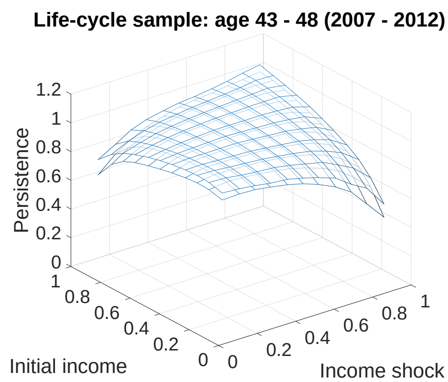


(f) 1964 cohort: 37-42

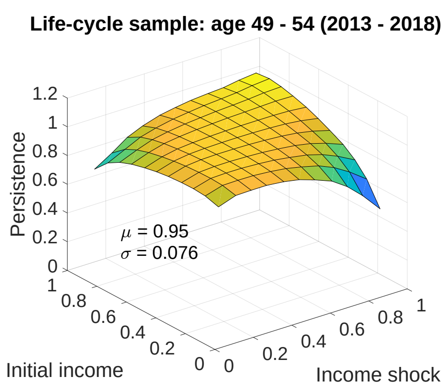
Figure B.5: Income persistence of y_t for 1964 birth cohort



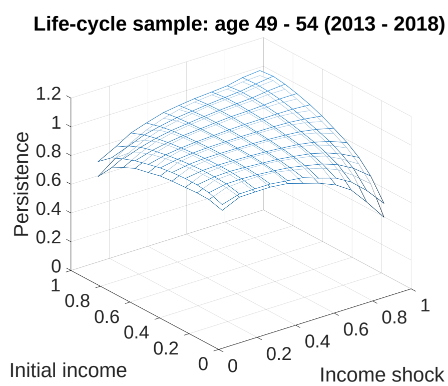
(a) 1964 cohort: 43-48



(b) 1964 cohort: 43-48



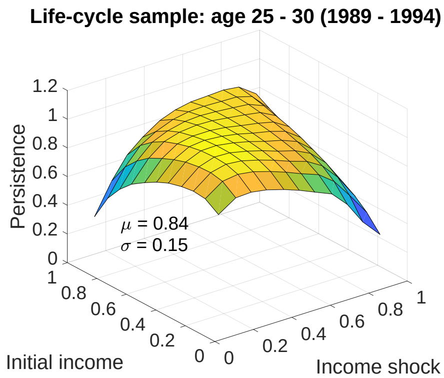
(c) 1964 cohort: 49-54



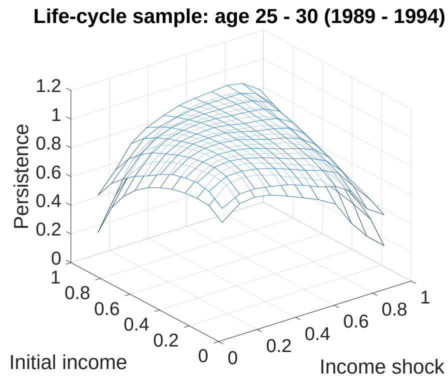
(d) 1964 cohort: 49-54

Figure B.6: Income persistence of y_t for 1964 birth cohort

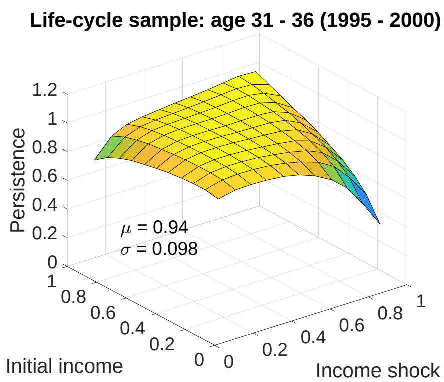
B.4 Life-cycle sample: persistent income



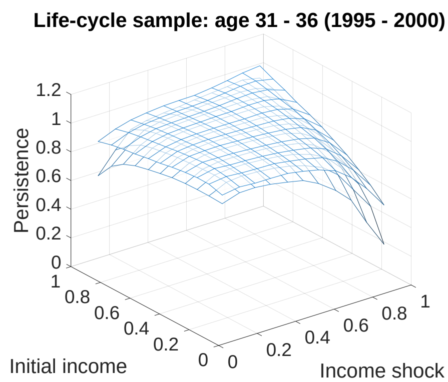
(a) 1964 cohort: 25-30



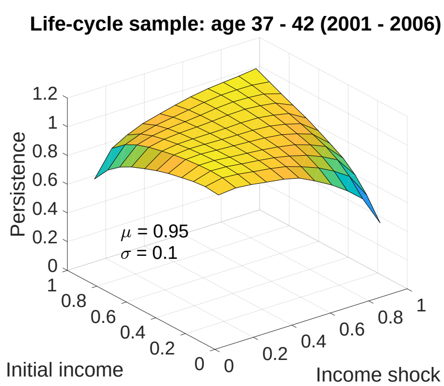
(b) 1964 cohort: 25-30



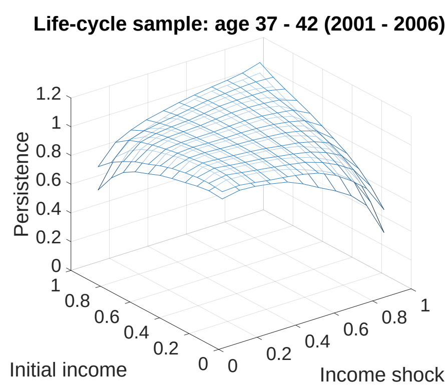
(c) 1964 cohort: 31-36



(d) 1964 cohort: 31-36

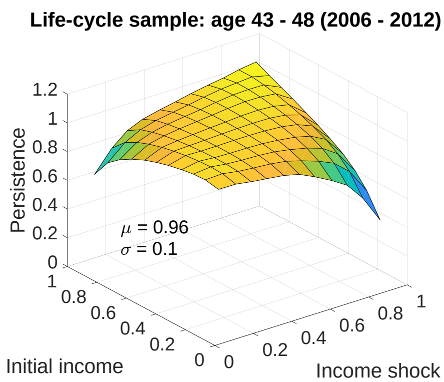


(e) 1964 cohort: 37-42

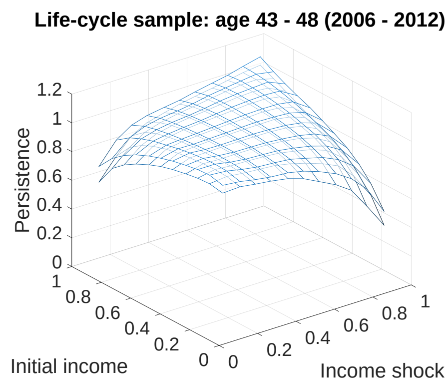


(f) 1964 cohort: 37-42

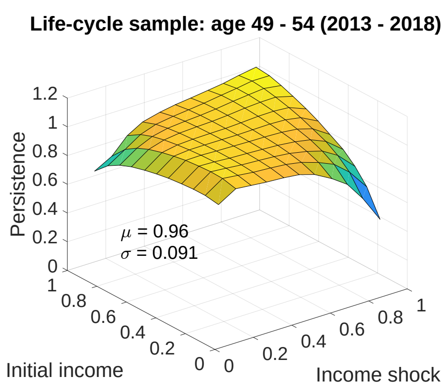
Figure B.7: Income persistence of η_t for 1964 birth cohort



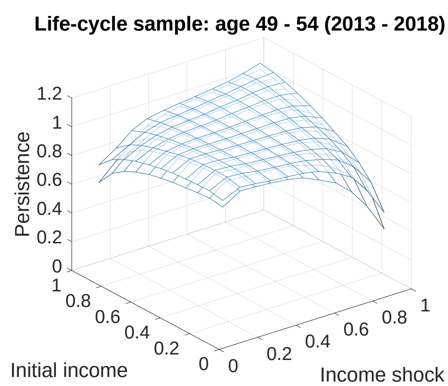
(a) 1964 cohort: 43-48



(b) 1964 cohort: 43-48



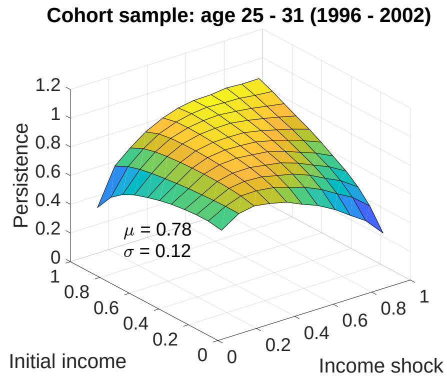
(c) 1964 cohort: 49-54



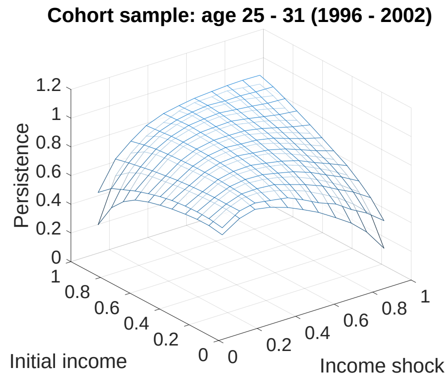
(d) 1964 cohort: 49-54

Figure B.8: Income persistence of η_t for 1964 birth cohort

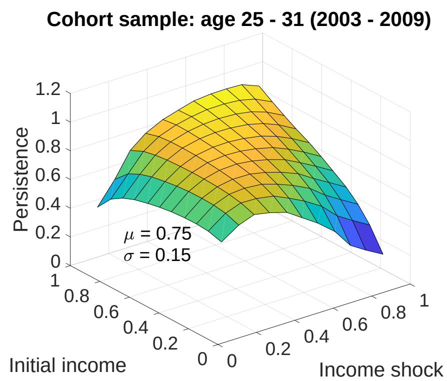
B.5 Cohort sample: residual log income



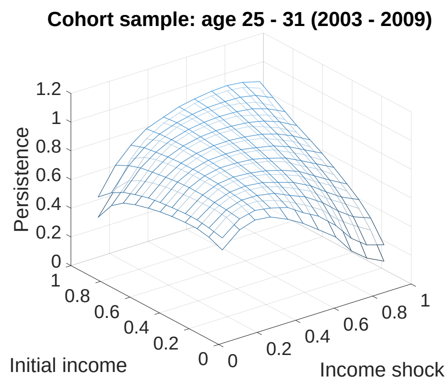
(a) 1971 cohort



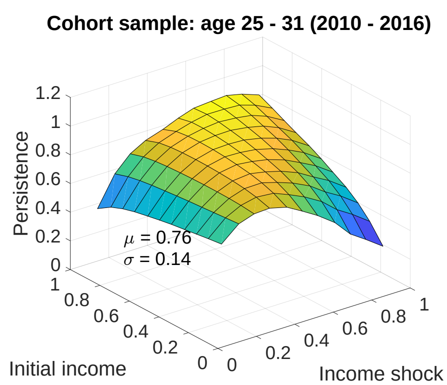
(b) 1971 cohort



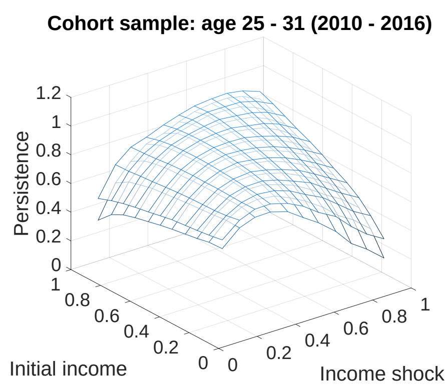
(c) 1978 cohort



(d) 1978 cohort

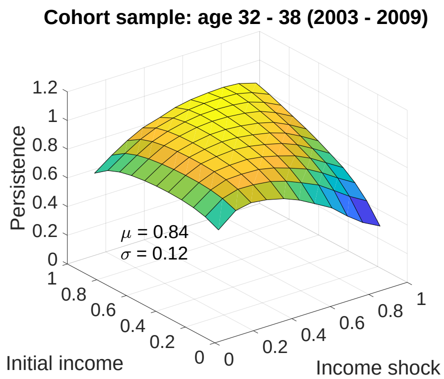


(e) 1985 cohort

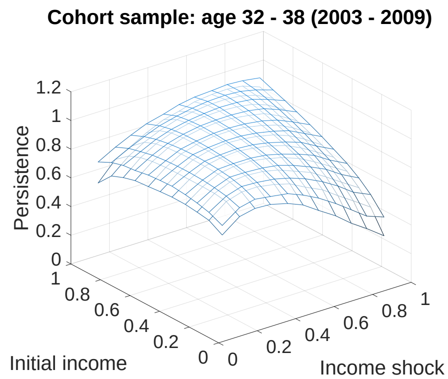


(f) 1985 cohort

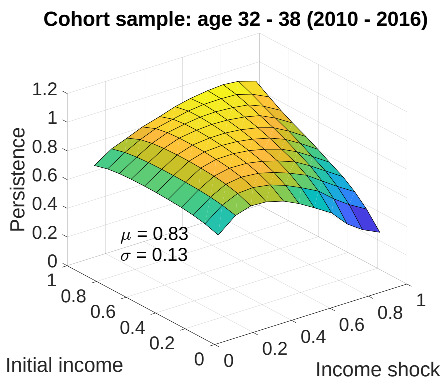
Figure B.9: Income persistence of y_t by birth year: mid 20s to early 30s



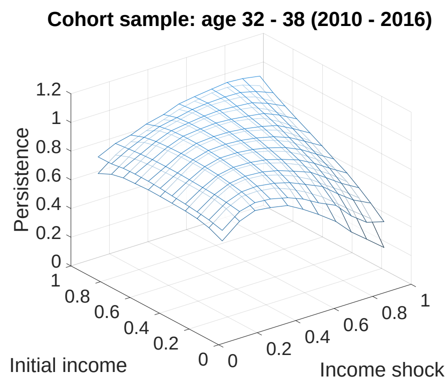
(a) 1971 cohort



(b) 1971 cohort

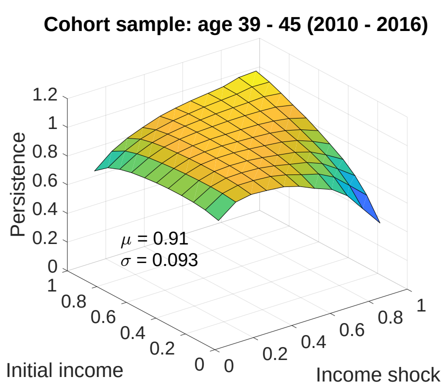


(c) 1978 cohort

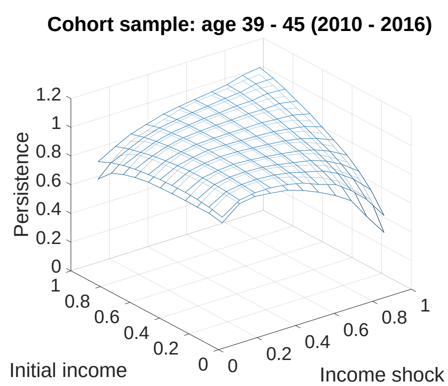


(d) 1978 cohort

Figure B.10: Income persistence of y_t by birth year: early to late 30s



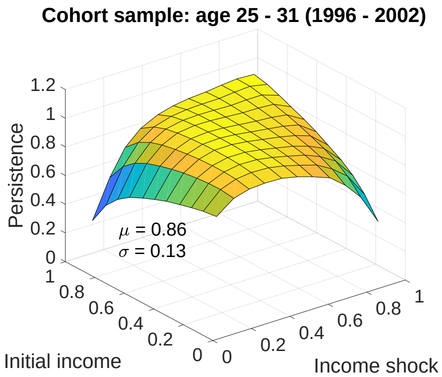
(a) 1971 cohort



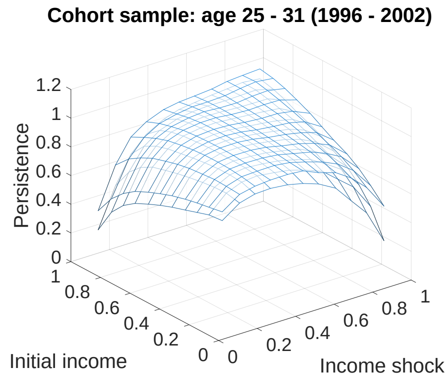
(b) 1971 cohort

Figure B.11: Income persistence of y_t by birth year: late 30s to mid 40s

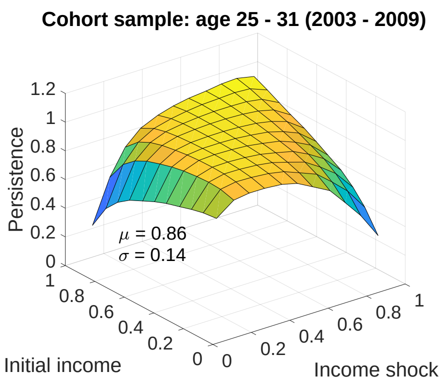
B.6 Cohort sample: persistent income



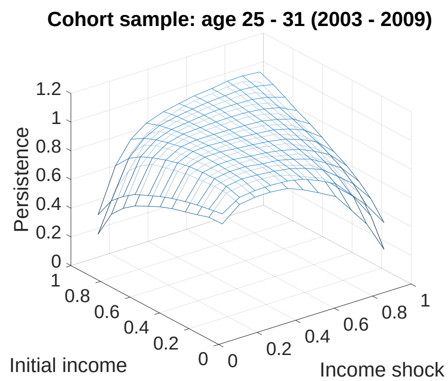
(a) 1971 cohort



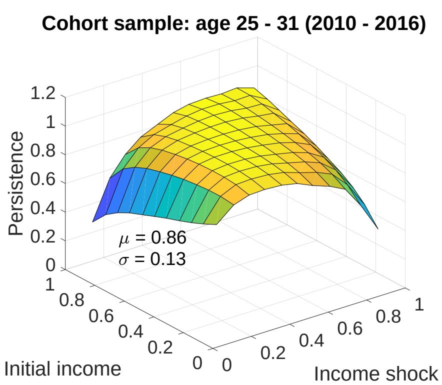
(b) 1971 cohort



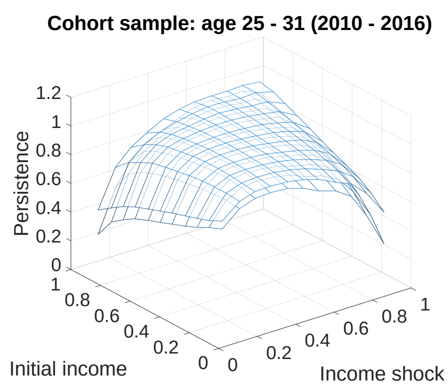
(c) 1978 cohort



(d) 1978 cohort

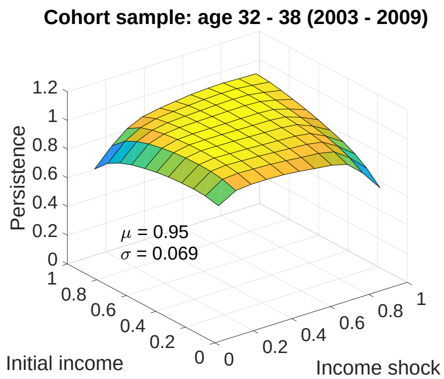


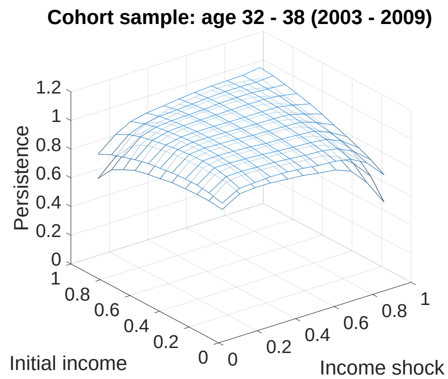
(e) 1985 cohort



(f) 1985 cohort

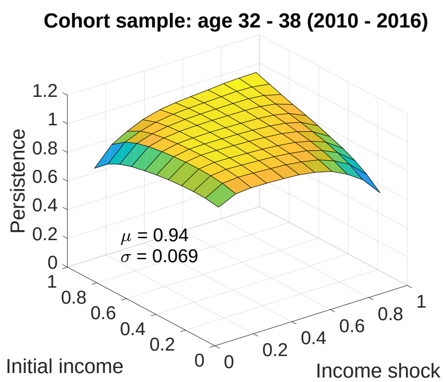
Figure B.12: Income persistence of η_t by birth year: mid 20s to early 30s

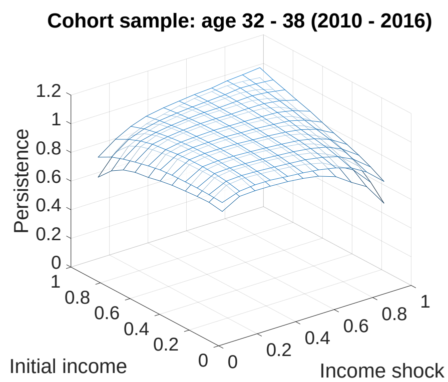




(a) 1971 cohort

(b) 1971 cohort

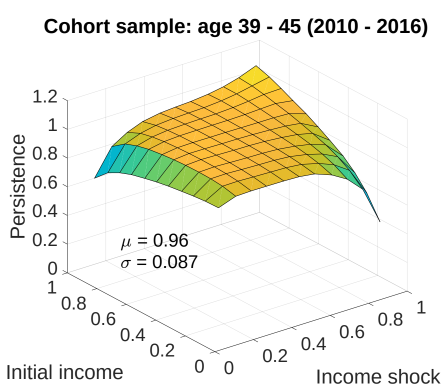


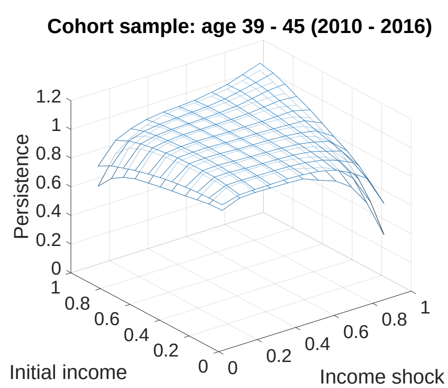


(c) 1978 cohort

(d) 1978 cohort

Figure B.13: Income persistence of η_t by birth year: early to late 30s





(a) 1971 cohort

(b) 1971 cohort

Figure B.14: Income persistence of η_t by birth year: late 30s to mid 40s