Precautionary savings motive in a volatile economy: Evidence from Argentine Administrative Data *

Juan Zurita[†] Emiliano Carlevaro[‡]

University of Edinburgh Adelaide University

Preliminary draft

31st August 2026

Abstract

The precautionary saving hypothesis implies that consumers save to protect themselves against labour income volatility. Evidence on this mechanism in emerging economies is scant and is often limited to aggregate cross-country data. We study the effect of labour income uncertainty on saving using a long panel of individual-level salary data from Argentina. We extend a standard model of individual labour income to accommodate multiple sectors of the economy. Using this model and the salary data, we estimate labour income volatility across sectors. We then estimate its effect on savings using district-level data on bank deposits and a spatial dynamic panel model. We find that labour income volatility has a modest positive effect on savings. The spatial model indicates that most of this effect arises from spatial spillovers from nearby districts.

*We thank Michelle Yu for her contributions to the project as a research assistant, as well as colleagues and conference participants for their valuable comments.

[†]Early Career Researcher, School of Economics. Corresponding author: juan.zurita@ed.ac.uk.

[‡]Lecturer, School of Economics.

1 INTRODUCTION

Understanding the structure and sources of earnings risk is central to macroeconomic and labour market research. A large literature, most notably Storesletten, Telmer and Yaron (2004), emphasises the importance of idiosyncratic income fluctuations for households' consumption, savings behaviour, and welfare. Their framework highlights how persistent and countercyclical earnings shocks shape precautionary savings and amplify aggregate fluctuations. While this foundational work has generated significant empirical and theoretical advances, most existing evidence is drawn from high-income economies, leaving far less known about the nature of earnings risk in emerging economies. This gap is particularly consequential given the greater macroeconomic volatility, higher informality, and more fragile social insurance systems that characterise these economies.

This study addresses this gap by extending the Storesletten, Telmer and Yaron (2004) approach to a large developing economy. We leverage a unique, high-frequency employer-employee matched dataset from Argentina's social security system, the Sistema Integrado Previsional Argentino (SIPA). The monthly panel is a representative and balanced sample of private formal employment relationships spanning 1996 to 2021, from which we use the 2011–2023 window, providing exceptionally detailed worker-level earnings histories across multiple economic cycles, recurrent macroeconomic crises, and significant structural changes in the Argentine labour market.

Using this administrative dataset, we estimate the persistence and cyclicality of individual earnings across major economic sectors. By disaggregating earnings risk across sectors, the analysis sheds light on how exposure to income fluctuations varies across industries, an essential dimension for understanding heterogeneity in consumption smoothing, precautionary savings, job mobility, and welfare outcomes. Sector-specific measures of earnings volatility are also crucial for calibrating macro-labour models that incorporate heterogeneous risk profiles and for evaluating the effectiveness of social insurance and labour market policies in environments marked by instability.

We then examine the precautionary saving motive by studying the effect of labour income volatility on savings during the period 2011 and 2023. Although individual-level data on savings are unavailable, we obtain district-level data on bank deposits, the wage bill, and the sectoral composition of employment. These data allow us to measure each district's exposure to different sectors of the economy and, consequently, to the underlying risks driving labour income volatility. Finally, we estimate a savings equation using a spatial dynamic panel data model with fixed effects using methods proposed by Lee and Yu (2010) and implemented in Simonovska (2025). The spatial model captures the endogenous effect of saving in one district on saving in nearby districts, and the spatial correlation in the wage bill, due to both input-output linkages between nearby districts and workers who commute to work but deposit their savings in their district of residence.

We find substantial dispersion in labour income volatility across sectors in the economy. We find statistically significant evidence that higher labour income volatility in a district increases savings in that district. The spatial model allows the decomposition of the effect of income volatility on savings into a direct effect, where a change in a district's income volatility affects only that district, and a spatial spillover effect, where a district is affected by changes in volatility in a nearby district. We indeed find that the stronger effect stems from this second spatial spillover.

Our findings help fill the substantial empirical gap regarding earnings risk in developing economies. By integrating the conceptual framework of Storesletten, Telmer and Yaron (2004) with granular Argentine administrative data, this research provides new insights into the nature of earnings dynamics in a volatile emerging market and their implications for workers, firms, and policymakers.

Our first contribution is to extend the model of Storesletten, Telmer and Yaron (2004) to multiple sectors. Our second contribution is that we find that differences in sectoral labour income risk affect saving behaviour. This finding is consistent with the precautionary saving motive, although the estimated semi-elasticity is economically small, at least on average.

2 LITERATURE REVIEW

A substantial body of research has examined the nature and implications of idiosyncratic earnings risk, particularly in the context of advanced economies. Early empirical studies documented substantial heterogeneity in individual income dynamics, highlighting the presence of transitory and persistent components in wage and earnings fluctuations (e.g., (Abowd and Card 1986); (MaCurdy 1982); (Meghir and Pistaferri 2004); (Pistaferri, Low and Meghir 2004)). Building on these insights, macroeconomic models began to incorporate rich income processes to understand household behaviour under incomplete markets.

Among the most influential contributions, Storesletten, Telmer and Yaron (2004) show that persistent idiosyncratic income shocks that are countercyclical play a pivotal role in shaping consumption, savings behaviour, and welfare. Their framework demonstrates how increased earnings volatility during downturns strengthens precautionary savings motives and amplifies aggregate fluctuations. Subsequent work extended and tested these insights, emphasizing the relevance of earnings shocks for explaining wealth inequality, portfolio choices, and consumption insurance (e.g., (Guvenen 2009); (Heathcote, Storesletten and Violante 2009); (Blundell 2017)). These studies consistently underscore persistent earnings risk as a key determinant of life-cycle behaviour and macroeconomic outcomes.

A parallel empirical literature has sought to measure the cyclicality and persistence of earnings risk more directly. Using panel data from the U.S. and other high-income economies, researchers provide evidence that both the variance and persistence of income shocks vary over the business cycle, increase during recessions, and differ by occupation, industry, and demographic group (e.g., (Carroll and Samwick 1997); (Hryshko 2012); (Arellano, Blundell and Bonhomme 2017)). More recent studies take advantage of administrative datasets to obtain cleaner measures of earnings dynamics, including social security or tax records ((Guvenen et al. 2015); (DeBacker et al. 2018)). Administrative data allow researchers to overcome limitations of survey-based earnings measures-notably measurement error, attrition, and top-coding-thereby producing more precise estimates of income risk.

Despite these advances, evidence from developing economies remains limited. Labour markets in these contexts are widely documented to exhibit higher volatility, weaker formal insurance mechanisms, and greater exposure to aggregate and sector-specific shocks. Informality is the central margin: Busso, Fazio and Levy (2012) show that a large informal sector in Mexico coexists with substantial misallocation and limited access to formal protection, while González-Rozada, Ronconi and Ruffo (2011) document for Argentina how thin unemployment insurance coverage and the prevalence of severance-based protection leave most workers exposed to job-loss risk. Macroeconomic downturns in the region also transmit directly to household welfare: Cruces, Glüzmann and López-Calva (2012) show that Argentina's recessions produced measurable deteriorations in infant health and schooling outcomes, evidence that households are unable to fully smooth aggregate shocks. However, much of this work relies on cross-sectional or low-frequency survey data, which constrain the ability to estimate persistent components of earnings risk, disentangle transitory from permanent shocks, or examine the cyclicality of risk at the individual level.

A more recent literature has begun to exploit administrative records to characterise income processes in Latin America directly. Most closely related to our work, Blanco et al. (2022) use Argentine social security records to document the evolution of the earnings distribution through the country's repeated macroeconomic crises, and Puggioni et al. (2022) provide comparable evidence on income dynamics and worker transitions in Mexico. Administrative data have also been used in the region for related purposes—for example Amarante et al. (2016) match Uruguayan social security records to vital statistics to evaluate a cash transfer programme. Relative to this work, our contribution is to estimate a structural earnings process rather than to describe distributional moments, and to carry the resulting risk estimates through to a behavioural outcome. The sectoral dimension of earnings risk—crucial in economies with strong heterogeneity in informality, productivity, and exposure to shocks—remains understudied.

Recent research on consumption smoothing uses individual-level data on consumption and savings. For example, Bergholt, Furlanetto and Mori (2026) study the Norwegian case and find that quantitatively as important as labour income uncertainty is consumption uncertainty, unexpected expenditures that need to be met like a health expenditure or the arrival of a newborn child.[1] We suspect results in an emerging economy will differ significantly for two reasons. First, the labour market is different, with the steady-state level of unemployment higher, a higher frequency of going into unemployment and smaller social protection systems. Second, smaller financial markets imply consumers have less scope for meeting unexpected consumption expenditure using credit.

Our study contributes to this growing literature in three aspects. First, we extend the Storesletten–Telmer–Yaron framework to accommodate sectoral differences, estimating the structure of earnings risk in a large developing economy using rich, high-frequency administrative data. Second, we provide sector-disaggregated estimates of the persistence and variance of individual earnings shocks, together with the cross-sector covariance structure of persistent innovations, offering novel evidence on heterogeneity in exposure to risk across sectors. Finally, we contribute to understanding the precautionary savings motive as a channel through which consumers protect themselves against labour income risk in an emerging economy.

This last aspect is becoming more important with the adoption of artificial intelligence, which is expected to raise labour demand in some sectors while contracting it in others, generating precisely the kind of heterogeneous, sector-specific income uncertainty we measure here (Acemoglu and Restrepo 2020; Acemoglu 2025). How households adjust their saving in response to differences in sectoral risk exposure is therefore of interest well beyond the Argentine case.

[1] The authors also consider wedding expenses as transitory expenditure shocks. We prefer to err on the side of rational expectation and consider them instead to be expected.

3 EMPIRICAL MODELS

3.1 Labour income dynamics

Following Storesletten, Telmer and Yaron (2004), adapted to monthly data, we model individual log wages $y^h_{i,t}$ as a the sum of two component: the first component captures the systematic component of wages that depends among other things on the age of the individual and aggregate income, the second component is a stochastic term that is idionsycratic to the individual such that

$$y^h_{i,t} = g(x^h_{i,t}, Y_t) + u^h_{i,t} \tag{1}$$

with the idiosyncratic component given by

$$u_{i,t} = \alpha_i + z_{i,t} + \varepsilon_{i,t}, \tag{2}$$

where

- $\alpha_i \sim \mathcal{N}(0, \sigma^2_\alpha)$ is an individual-specific fixed effect,
- $z_{i,t}$ is a persistent AR(1) component with state-dependent volatility,
- $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_\varepsilon)$ is an i.i.d. transitory shock.

The persistent component follows

$$z_{i,t} = \rho\, z_{i,t-1} + \eta_{i,t}, \qquad \eta_{i,t} \sim \mathcal{N}\Big(0, \sigma^2_{z,s}\Big),$$

where the innovation variance depends on the aggregate state $s \in \{\text{good}, \text{bad}\}$:

$$\sigma^2_{z,s} = \begin{cases} \sigma^2_{z,\text{good}} & \text{if } s = \text{good}, \\ \sigma^2_{z,\text{bad}} & \text{if } s = \text{bad}. \end{cases}$$

The probability of being in the bad state in any given month is

$$\pi \equiv \Pr(s = \text{bad}).$$

Hence, the parameter vector to be estimated is

$$\theta = \Big(\rho,\ \sigma_\alpha,\ \sigma_\varepsilon,\ \sigma_{z,\text{good}},\ \sigma_{z,\text{bad}},\ \pi\Big).$$

Estimation

The model is estimated by matching cross-sectional moments of log earnings to their theoretical counterparts implied by θ. Because the panel is monthly, all moments below are stated at monthly frequency. For a worker of age h (in years) we let $m = 12h$ denote age in months, measured from labour market entry.

Averaging over business cycle states, the model-implied cross-sectional variance of y at age h is

$$\operatorname{Var}\Big(y_{i,t} \mid \text{age} = h\Big) = \sigma_\alpha^2 + \sigma_\varepsilon^2 + \underbrace{\mathbb{E}[\sigma_{z,s}^2] \sum_{j=0}^{m-1} \rho^{2j}}_{\text{accumulated persistent variance}}, \tag{3}$$

where

$$\mathbb{E}\Big[\sigma_{z,s}^2\Big] = (1-\pi)\,\sigma_{z,\text{good}}^2 + \pi\,\sigma_{z,\text{bad}}^2,$$

and the lag-1 (one-month) autocovariance at age h is

$$\operatorname{Cov}(y_{i,t}, y_{i,t-1} \mid \text{age} = h) = \sigma_\alpha^2 + \rho\,\mathbb{E}\Big[\sigma_{z,s}^2\Big] \sum_{j=0}^{m-2} \rho^{2j}, \qquad m > 1. \tag{4}$$

Two points about (4) deserve emphasis, since both bear directly on identification. First, the summation runs to $m-2$ in *months*, matching (3); stating the variance in months and the autocovariance in years would make the two moments mutually inconsistent. Second, the fixed effect σ_α^2 enters the autocovariance, because α_i is common to $y_{i,t}$ and $y_{i,t-1}$ by construction. Omitting it would load the entire level of the autocovariance profile onto ρ and $\mathbb{E}[\sigma_{z,s}^2]$, biasing the estimated persistence.

Note also that for $|\rho| < 1$ the geometric sums in (3) and (4) converge to $(1-\rho^2)^{-1}$ and $\rho(1-\rho^2)^{-1}$ respectively, and do so after approximately $\log(0.01)/(2\log\rho)$ months. Unless the monthly ρ is very close to unity, the model-implied moments are therefore effectively flat in age over the range observed in the data, and the age profile carries little identifying information. Section 5.1 returns to this point.

The model is indexed by the six parameters (monthly frequency)

$$\theta \;=\; \Big(\rho,\ \sigma_\alpha,\ \sigma_\varepsilon,\ \sigma_{z,\text{good}},\ \sigma_{z,\text{bad}},\ \pi\Big).$$

Let $\widehat{m}_{\text{data}}(h)$ denote the empirical cross-sectional variance and lag-1 autocovariance for each age h (averaged across months/years or computed as a time series $m_t(h)$ and then averaged). Stack the empirical moments into the vector

$$\widehat{m}_{\text{data}} = \Big(\widehat{\text{Var}}(h_1), \ldots, \widehat{\text{Var}}(h_K),\ \widehat{\text{ACov}}_1(h_1), \ldots, \widehat{\text{ACov}}_1(h_K)\Big)',$$

and denote the model-implied moments by $m_{\text{model}}(\theta)$ with the same ordering.

GMM estimator (two-step). The GMM estimator minimizes a quadratic distance between empirical and model moments. The two-step procedure is:

1. **First step (initial weight).** Choose an initial positive definite weighting matrix W_0 (commonly I) and compute

$$\hat{\theta}^{(1)} = \arg\min_{\theta} \Big(\widehat{m}_{\text{data}} - m_{\text{model}}(\theta)\Big)' W_0 \Big(\widehat{m}_{\text{data}} - m_{\text{model}}(\theta)\Big).$$

2. **Compute sample covariance of moment residuals.** If m_t denotes the vector of monthly empirical moments at time t and T is the number of months used, form the sample covariance

$$\widehat{S} = \frac{1}{T-1}\sum_{t=1}^{T} \Big(m_t - m_{\text{model}}(\hat{\theta}^{(1)})\Big)\Big(m_t - m_{\text{model}}(\hat{\theta}^{(1)})\Big)'.$$

3. **Second step (optimal weight).** Use the optimal weighting matrix $W^* = \widehat{S}^{-1}$ (or a regularized inverse if needed) and compute

$$\hat{\theta}^{(2)} = \arg\min_{\theta} \Big(\widehat{m}_{\text{data}} - m_{\text{model}}(\theta)\Big)' W^* \Big(\widehat{m}_{\text{data}} - m_{\text{model}}(\theta)\Big).$$

Asymptotic variance. Let $G(\theta) = \partial m_{\text{model}}(\theta)/\partial\theta'$ denote the $M \times p$ Jacobian of model moments w.r.t. parameters (evaluated at $\hat{\theta}^{(2)}$), and $S = \widehat{S}$. The asymptotic covariance matrix of $\hat{\theta}^{(2)}$ is approximated by the sandwich formula

$$\widehat{\text{Var}}(\hat{\theta}^{(2)}) = \Big(G'W^*G\Big)^{-1} G'W^*SW^*G \Big(G'W^*G\Big)^{-1},$$

with $W^* = S^{-1}$ (so the middle term simplifies when W^* equals S^{-1}).

Estimates of θ for the full sample and by sector are reported in Section 5.1, Table 1.

3.2 Savings model

For each district, we compute the labour income in the district as $Y_d = n_{d,t} \times \bar{w}_{d,t}$ where $n_{d,t}$ is the number of employees and $\bar{w}_{d,t}$ the median salary. The share of sector k in labour income of district d at time t is $v_{d,t;k} \equiv y_{d,t;k}/Y_{d,t}$.

Let Σ_η be the variance-covariance matrix of labour income across sectors as estimated in the previous subsection and defined in equation (11) We assume that the variance of income across sectors is sufficiently stable over time to consider it fixed.

For a district, the variance of labour income is given by the exposition to the fundamental variance Σ_η such that the labour income variance of district d at time t denoted by $\sigma_{d,t}$ is

$$\sigma^2_{d,t} = \mathbf{v}'_{d,t}\, \mathbf{\Sigma}_\eta\, \mathbf{v}_{d,t}, \tag{5}$$

We study the ratio of savings to labour income. The dependent variable is $s_{d,t} = \Delta D_{d,t}/Y_{d,t}$, where the numerator is the first difference of the stock of real deposits in district d at time t and the denominator is the real wage bill.

Our model for savings is a dynamic spatial Durbin model. We focus on this spatial model for two reasons. First, economic activity and labour are spatially correlated, as an economic improvement in a district has spatial spillover in nearby districts. This implies that labour income is spatially correlated. Second, workers commute to nearby districts for work but may deposit their savings at the bank branch closest to their residential district. This implies that higher savings in a district is correlated with labour income in a nearby district.

The distance between districts is captured by the spatial weight matrix W. Our baseline result uses a queen contiguity matrix such that districts that share a border are linked together. On average, a district shares a border with 4.85 other districts. We row normalise the matrix for estimation.

Let $y_{d,t} \equiv \log(Y_{d,t})$ and $\boldsymbol{x}_{d,t} = \begin{bmatrix} y_{d,t} & y_{d,t-1} & \log \sigma^2_{d,t} \end{bmatrix}$ be a row vector for district d in time t then the vector of saving ratios for all districts at time t follows

$$\boldsymbol{s}_t = \phi \boldsymbol{s}_{t-1} + \rho W \boldsymbol{s}_t + X_t \boldsymbol{\beta} + W X_t \boldsymbol{\theta} + \boldsymbol{c}_i + a_t \boldsymbol{\iota}_n + \boldsymbol{\epsilon}_t \tag{6}$$

with ϕ the time autoregression parameter, ρ the spatial autocorrelation parameter, X_t a matrix that vertically stacks $\boldsymbol{x}_t$ for a given t, , $\boldsymbol{\theta}$ is a parameter vector of spatial lagged covariates, $\boldsymbol{c}_i$ is a vector of district fixed effect, a_t is a scalar time effect and $\boldsymbol{\iota}_n$ is an $n \times 1$ column vector of ones.

The parameter vector θ on the spatially lagged covariates WX_t captures how labour income in a district is shaped by the labour income of nearby districts. The 3×1 paremeter vector $\boldsymbol{\beta}$ contains the main parameters of interest: the response of savings to income, captured by the coefficient on $y_{d,t}$; the response to the lag of $y_{d,t}$; and the semi-elasticity to the variance of labour income.

The idiosyncratic error term $\boldsymbol{\epsilon}_t$ is independent and identically distributed across d and t with a zero mean and variance σ_ϵ^2. Estimation proceeds in a quasi-maximum-likelihood framework (QML), assuming normality of the idiosyncratic error and transforming the data using the group mean transformation to alleviate the asymptotic bias generated by the incidental parameter problem, as suggested by Lee and Yu (2010). We use the implementation of Lee and Yu (2010)' QML estimator in R by Simonovska (2025).

The marginal effect of a covariate in the model in (6) is heterogeneous across districts due to their different location in space. We report three common summary statistics. The direct effect of covariate x_k on savings is $\mathbb{E}[\partial s_i / \partial x_{i,k}]$, the average across districts of the marginal effect of a change in covariate x_k for district i on savings in district i. The direct effect, or own spillover effect, captures the effect of a change in, for example, the wage bill in district i including the feedback loop where the change in the wage bill in district i affects savings in district i, which then spillovers to other districts through the parameter ρ, and it feeds back to the focal district i. The indirect effect of covariate x_k on savings is $\mathbb{E}[\partial s_j / \partial x_{i,k}]$, which captures the average across districts of the spillover effect of a change in covariate x_k in district i on all *other* districts. For example, marginally increasing the wage bill in district i affects the savings of other districts through the parameters ρ and θ. The average of this spillover effect on savings across all districts represents the indirect effect.[2]

4 DATA

For our data on salaries our sample covers the period 2011 through 2023. To estimate the volatility of labour income, we rely on individual and monthly level data on salaries of employees in the private sector. Our subsequent analysis of savings relies on quarterly district-level data from 2014q1 to 2023q3. Our main results on savings are based on the period 2014q3 through 2023q3.

[2]This description corresponds to the average total impact *from* an observation as described in LeSage and Pace (2009, Ch. 1) see also **KelejianPiras2017** which label this effect as the emanating effect

4.1 Labour income

Our primary data source is an employer-employee matched panel derived from Argentina's social security administrative records, the Sistema Integrado Previsional Argentino (SIPA). These records originate from the monthly sworn statements that employers are legally required to submit to Argentina's tax authority, the Administracion Federal de Ingresos Publicos (AFIP). The dataset contains information on payroll subject to social security contributions, covering all formal employees in the private sector. We use a 3 percent random, anonymized subsample of private-sector employees spanning the period 1996-2021.

The dataset provides detailed information on employees' monthly gross labour earnings, encompassing all taxable forms of compensation, including base salary, overtime pay, performance and seasonal bonuses, paid vacations, paid sick leave, and severance payments. It also includes demographic characteristics such as gender, year of birth, and the province of the establishment. To preserve confidentiality, earnings above the 98th percentile within each industry are anonymized. However, the data do not include information on employees' education levels. Employer-level information includes the firm's four-digit industry classification, and both employees and employer-employee matches are assigned anonymized unique identifiers, enabling the tracking of individual workers and formal employment relationships over time.

The sample is representative of the formal private-sector workforce across industries, regions, and contract types (e.g., full-time, temporary, and internship positions). The number of observations grows from approximately 130,000 workers in 1996 to about 176,000 in 2021. Given that formal private employment accounts for roughly 30-40 percent of total employment in Argentina over this period (including independent and self-employed workers), the sample represents around 1 percent of the employed population in any given year.

4.2 Savings

We use data on deposits of the private sector as compiled by the Argentine Central Bank. This quarterly series contains, at the city level, the stock of bank deposits by the private sector. Sadly, this series does not distinguish between deposits by firms and by households. In our baseline analysis, we combine deposits in pesos and US dollars. We match city-level data to the district using a corresponding table from the Argentine Geographic Institute and manual matching in some cases.

We exclude the City of Buenos Aires from our savings analysis. The City, as the political capital and financial centre of the country, holds a disproportionately large share of deposits, mostly from private corporations headquartered in the city.

To compute the wage bill for each district, we use two databases from the Office of Labour Statistics. The first one contains the number of private-sector employees and the median salary per district for a given quarter. The second dataset includes the number of employees in a district by economic sector at the 2-digit level. We then aggregate these 2-digit level classifications to 1-digit level resulting in an economy with 15 sectors.

The data from the Office of Labour Statistics uses a modern economic activity classification scheme as designed by the statistical office known as CLANAE 2010 (revised version 2017). The salary data used to estimate income volatility, uses instead the classifier known as CLANAE 1997. To map our wage bill data to the volatility of labour income, we use data from the Statistical Office INDEC to classify most 2-digit-level codes to the corresponding CLANAE 1997 sector-level (1-digit) classification. In the few cases where the mapping is not unique, we based our classification on the number of subsectors at the 2-digit level that most closely match the corresponding aggregated 1-digit classification in the CLANAE 1997.

We source district boundaries from the *Instituto Geografico Nacional*.

5 PRELIMINARY RESULTS

5.1 Labour income uncertainty

We first report estimates of the earnings process, sector by sector, obtained from individual-level salary data. Table 1 presents the two-step GMM estimates of θ; Figures 1–3 plot the empirical variance and lag-1 autocovariance profiles against their model-implied counterparts, with the remaining sectors shown in Appendix 7.1.

Table 1: Baseline earnings process: two-step GMM estimates by sector

Sample	ρ	σ_α	σ_ε	$\sigma_{z,\text{good}}$	$\sigma_{z,\text{bad}}$	π
Manufacturing	0.763	0.121	0.108	0.528	0.000	0.034
Transport	0.873	0.261	0.219	0.453	0.399	0.344
Real Estate Services	0.898	0.323	0.259	0.353	0.470	0.384
Hotels and Catering	0.688	0.101	0.100	0.811	0.000	0.492
Agriculture	0.949	0.299	0.339	0.234	0.000	0.040
Trade	0.890	0.244	0.237	0.343	0.000	0.051
Memorandum: implied annual persistence ρ^{12}						
Manufacturing 0.039 Transport 0.196 Real Estate 0.275 Hotels 0.011 Agriculture 0.534 Trade 0.247						

Notes: Parameters are estimated at monthly frequency by two-step GMM, matching cross-sectional variance and lag-1 autocovariance profiles by age over $h \in [25, 60]$, separately for each sector. π is estimated rather than imposed. The memorandum item converts the estimated monthly persistence to an annual equivalent for comparison with the literature, which typically reports $\rho_{\text{annual}} \approx 0.95$–$0.99$.

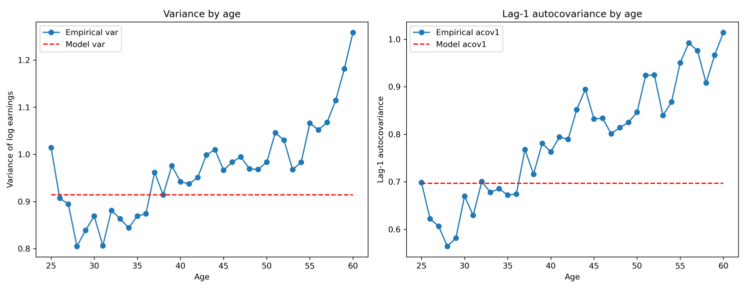
Three features of Table 1 require comment, and we flag them openly because they bear on how much weight the sectoral risk estimates can carry.

First, the estimated monthly persistence implies annual persistence far below the values conventionally found in the earnings dynamics literature. A monthly ρ of 0.873 in Transport corresponds to an annual persistence of $0.873^{12} \approx 0.20$, against the 0.95–0.99 range reported by Storesletten, Telmer and Yaron (2004) and subsequent work. Matching an annual persistence of 0.95 would require a monthly ρ of approximately 0.996.

Second, and as a direct consequence, the model-implied moment profiles are approximately flat in age. With $\rho = 0.873$, the geometric sum in (3) converges to within one per cent of its limit after roughly 17 months, so by the youngest age observed in the data ($h = 25$, i.e. $m = 300$) the persistent component has long since reached its stationary variance. The model therefore predicts a constant cross-sectional variance of log earnings over the entire age range, whereas the empirical profiles rise substantially with age. The life-cycle fanning-out of the earnings distribution, which is the feature that identifies ρ and σ_α in the Storesletten–Telmer–Yaron framework, is not reproduced by the estimated model.

Third, the estimated $\sigma_{z,\text{bad}}$ lies at the boundary value of zero in four of the six sectors, and in those sectors is therefore *below* $\sigma_{z,\text{good}}$. Taken literally this implies procyclical rather than countercyclical idiosyncratic risk, the opposite of the pattern that motivates the framework. We regard this as more likely to reflect an identification problem than an economic finding: the two states enter the moment conditions only through the mixture $\mathbb{E}[\sigma^2_{z,s}] = (1-\pi)\sigma^2_{z,\text{good}} + \pi\sigma^2_{z,\text{bad}}$, so without a restriction tying the state s to the observed aggregate cycle, the labels "good" and "bad" are not separately identified and the optimiser is free to allocate the mixture to whichever state it likes. [3]

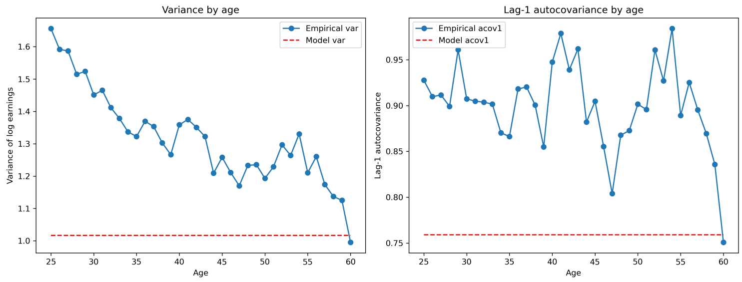
Figure 1: Variance and Lag-1 Autocovariance of Log Earnings, the Transport Sector



Notes: The left panel plots the cross-sectional variance of log earnings by age; the right panel plots the lag-1 (one-month) autocovariance by age. Solid blue lines with markers are empirical moments computed from the SIPA sample; red dashed lines are the model-implied moments evaluated at the two-step GMM estimates for Transport reported in Table 1. Ages 25–60. The model-implied profiles are flat because the estimated monthly ρ implies that the persistent component reaches its stationary variance well before age 25; see the discussion in Section 5.1.

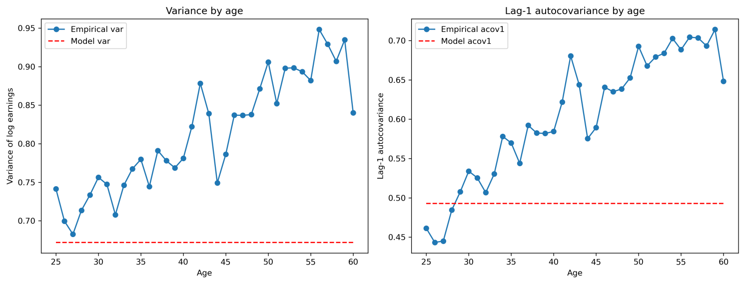
[3]These three issues are addressed in the multi-sector specification of Appendix 7.2, which imposes a common ρ across sectors and replaces the two-state structure with a full covariance matrix. Pending their resolution, the two-state results in this table should be read as preliminary.

Figure 2: Variance and Lag-1 Autocovariance of Log Earnings, the Real Estate Services Sector



Notes: The left panel plots the cross-sectional variance of log earnings by age; the right panel plots the lag-1 (one-month) autocovariance by age. Solid blue lines with markers are empirical moments computed from the SIPA sample; red dashed lines are the model-implied moments evaluated at the two-step GMM estimates for Real Estate Services reported in Table 1. Ages 25–60. The model-implied profiles are flat because the estimated monthly ρ implies that the persistent component reaches its stationary variance well before age 25; see the discussion in Section 5.1.

Figure 3: Variance and Lag-1 Autocovariance of Log Earnings, the Manufacturing Sector



Notes: The left panel plots the cross-sectional variance of log earnings by age; the right panel plots the lag-1 (one-month) autocovariance by age. Solid blue lines with markers are empirical moments computed from the SIPA sample; red dashed lines are the model-implied moments evaluated at the two-step GMM estimates for Manufacturing reported in Table 1. Ages 25–60. The model-implied profiles are flat because the estimated monthly ρ implies that the persistent component reaches its stationary variance well before age 25; see the discussion in Section 5.1.

5.2 Savings equation

We now turn to the role of labour income uncertainty in driving private savings in districts.

Endogenous spatial dependence, captured by the spatial autoregressive parameter ρ, is statistically significant and positive, albeit modest in magnitude. This confirms a positive spatial spillover in savings, whereby higher savings in a given district positively influence savings rates in neighboring areas.

We also find evidence of time dynamics in savings, with the estimated coefficient on the lagged savings ratio being negative and statistically significant at -0.121. This contrasts with the positive autoregressive coefficient (on the order of 0.60) reported by Grigoli, Herman and Schmidt-Hebbel (2018), who examine cross-country variation in savings-to-GDP ratios using annual data.

Both coefficient estimates for labor income volatility remain positive and statistically signi-

Table 2: **STATIC AND DYNAMIC SPATIAL DURBIN MODEL ESTIMATES**. The dependent variable is the savings ratio defined as the first difference of private deposits to the private sector wage bill. All variables deflated by CPI.

Explanatory variable	**SDM Ind.**	**SDM Two-way**	**Dynamic SDM**
ρ	0.32***	0.06***	0.06***
	(0.01)	(0.01)	(0.01)
Saving ratio$_{d,t-1}$			−0.12***
			(0.01)
log(Labour income$_{d,t}$)	0.22***	0.22***	0.22***
	(0.03)	(0.03)	(0.03)
$\log(\sigma^2_{d,t})$	0.06***	0.08***	0.08***
	(0.02)	(0.02)	(0.01)
Neighbours' log(Labour income$_{d,t}$)	0.02	0.02	0.01
	(0.04)	(0.06)	(0.06)
Neighbours' log(Labour income$_{d,t-1}$)	−0.02	−0.10*	−0.12*
	(0.04)	(0.06)	(0.06)
Neighbours' $\log(\sigma^2_{d,t})$	0.01	0.12***	0.12***
	(0.03)	(0.03)	(0.03)
Log likelihood	−876.84	198.71	
R^2	0.16	0.25	
Adjusted R^2	0.07	0.01	

Notes: Standard errors are reported in parentheses. Observations are 15,960 with 420 districts and 39 time periods.
$^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

ficant, aligning with the baseline static non-spatial models.

Table 3 presents the marginal effects derived from the dynamic spatial Durbin model. Panel A reports short-run impacts, which do not account for the temporal autoregressive dynamics of the savings ratio. The "Direct" column measures the average response of a district's savings ratio to a shift in its own covariate. The "Indirect" column captures spatial spillovers, the average cumulative impact that a covariate shift in one district exerts on all other districts. Finally, the "Total" column provides the overall effect across the spatial system.

Panel B of Table 3 displays the long-run steady-state marginal effects, which constitute our preferred estimates given the dynamic structure of the empirical specification.[4]

Regarding local economic drivers, a 10% increase in average labor income raises the district savings ratio by approximately 0.025.

The parameter estimate for labor income uncertainty remains remarkably stable across specifications. As shown in the final column of Table 3, the estimated total semi-elasticity of savings with respect to labor income uncertainty is 0.00202, implying that a 10% increase in

[4]Due to the negative temporal persistence parameter ($\hat{\phi} = -0.121$), the long-run marginal effects are slightly smaller than their short-run counterparts.

Table 3: **MARGINAL EFFECTS**. Dynamic Spatial Durbin model.

Explanatory variable	Direct	Indirect	Total
	A. Short-term		
log(Labour income$_{d,t}$)	0.219***	0.027	0.246***
	(0.034)	(0.061)	(0.065)
log(Labour income$_{d,t-1}$)	−0.008	−0.136**	−0.144**
	(0.035)	(0.063)	(0.067)
log($\sigma^2_{d,t}$)	0.085***	0.143***	0.228***
	(0.017)	(0.031)	(0.034)
	B. Long-term		
log(Labour income$_{d,t}$)	0.195***	0.023	0.218***
	(0.030)	(0.054)	(0.057)
log(Labour income$_{d,t-1}$)	−0.007	−0.121**	−0.128**
	(0.031)	(0.056)	(0.060)
log($\sigma^2_{d,t}$)	0.076***	0.126***	0.202***
	(0.015)	(0.028)	(0.030)

Notes: Standard errors are reported in parentheses. Observations are 15,960 with 420 districts and 39 time periods.
$^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

labor income uncertainty elevates the savings ratio by roughly 0.02.

Crucially, decomposing this total effect reveals that spatial spillovers dominate the total impact: an estimated 60% of the response originates from neighboring areas. On average, a 10% increase in labor income uncertainty in neighboring districts raises the savings ratio of the focal district by 0.0126. As posited above, these prominent spatial spillovers likely reflect inter-district commuting flows and local input-output linkages that transmit labor market shocks across administrative borders.

5.3 Robustness Analysis

5.3.1 Alternative periods

We consider three different subsample periods. The first one is 2016q1-2018q1, which is a "tranquil" period with minor variation in aggregate variables like the exchange rate. The second one is before 2019q3, running from 2014q1-2019q2, avoiding the political uncertainty generated by the results of the primary national election in August 2019, which resulted in significant macroeconomic volatility, particularly in the exchange rate. The third subsample runs from 2019q3 till the end of our data in 2023q3, covering both the change in the federal government since 2019 and the covid period, which we may label a "high volatility" period.

We present the marginal effects of the 3 subsamples from the dynamic Spatial Durbin model in Table 4. From the Panel A of the tranquil period of 2016-2018 estimated marginal effects are smaller in magnitued than the baseline ones. Focusing on statistically significant effects, the direction of the effects is the same as in the baseline case. As in the full-sample case, the effect of labour income on savings comes primarily through the direct effect, with a statistically non-significant estimated indirect effect, as in the baseline case. The indirect effect of income volatility of 0.14 is positive and is larger than the direct effect, replicating the pattern in the baseline results.

Table 4: **MARGINAL EFFECTS IN ALTERNATIVE PERIODS**. Long-term effects from the dynamic spatial Durbin model.The dependent variable is the savings ratio defined as the first difference of private deposits to the private sector wage bill. All variables deflated by CPI.

Explanatory variable	**Direct**	**Indirect**	**Total**
A. 2016Q1–2018Q1			
log(Labour income$_{d,t}$)	0.20***	−0.06	0.15
	(0.07)	(0.13)	(0.13)
log($\sigma^2_{d,t}$)	−0.03	0.14*	0.10
	(0.04)	(0.07)	(0.07)
B.Before 2019Q3			
log(Labour income$_{d,t}$)	0.16***	−0.02	0.13**
	(0.03)	(0.06)	(0.06)
log($\sigma^2_{d,t}$)	0.03*	0.13***	0.16***
	(0.02)	(0.03)	(0.04)
C. Since 2019Q3			
log(Labour income$_{d,t}$)	0.29***	0.10	0.39***
	(0.06)	(0.10)	(0.11)
log($\sigma^2_{d,t}$)	0.15***	0.14**	0.29***
	(0.03)	(0.06)	(0.06)

Notes: Standard errors are reported in parentheses. Observations are 3780, 9240 and 6720 for each period.
$^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Statistical significance improves with larger, more volatile remaining periods. The period that covers 2014q3 through 2019q3 shows effects that are similar to the full-sample, although the magnitude of all effects is slightly smaller. The last period of "high volatility" from 2019q3 to 2023q3, results in estimated larger effects across all variables. Interestingly, in this last period, the direct effect of income volatility is larger than the indirect effect. The total effect of income volatility is almost twice that of the previous period before 2019q3. This last period, is characterised by a drawdown of savings rather than a buildup. The average savings ratio is 0.061 and 0.040 for the first two periods, and -0.032 for the subsample since 2019q3. Given the larger magnitude of this effect relative to other periods, it may suggest an asymmetric effect of income volatility on savings.

5.4 Limitations

Some limitations qualify these findings. First, while district fixed effects absorb static demographic differences, our savings equation does not explicitly control for time-varying population age structures—a potential oversight given that the life-cycle hypothesis predicts systematic age-driven variation in saving rates. Second, treating $\mathbf{\Sigma}_\eta$ as time-invariant assumes that sectoral risk profiles remained static despite multiple macroeconomic regime shifts, attributing all temporal variation in $\sigma^2_{d,t}$ in (5) solely to shifts in local sectoral composition. Third, our formal bank deposit data aggregates firm and household holdings and excludes informal precautionary assets. Because Argentine households frequently hold foreign currency outside the domestic banking system during periods of heightened risk, our findings likely represent a lower bound on the true precautionary response. Finally, because deposits are recorded at the bank branch level rather than by depositor residence, commuting patterns introduce spatial measurement error that our spatial Durbin specification mitigates but does not fully eliminate.

6 SOME REMARKS

Using an individual-level labour income model, we estimate sectoral income volatility in Argentina between 2011 and 2023. We then evaluate the impact of this volatility on district-level savings using a spatially dynamic Durbin model with fixed effects.

Our empirical findings reveal substantial dispersion in labour income volatility across sectors. Crucially, higher labour income volatility within a district statistically increases local savings rates. A significant portion of this effect operates through spatial spillovers: shifts in income risk within neighbouring districts directly elevate the savings ratio in the focal district.

Two methodological caveats limit the scope of these results and warrant plain statement. First, our sectoral estimates of the earnings process do not reproduce the life-cycle widening of the earnings distribution typically used to identify these models in advanced-economy literature. Second, the two-state cyclical structure is not separately identified without an external dating of the aggregate business cycle. Resolving both issues-by anchoring the aggregate state to an independent business cycle chronology and allowing persistence closer to the unit root found in related literature-remains the primary priority for future research.

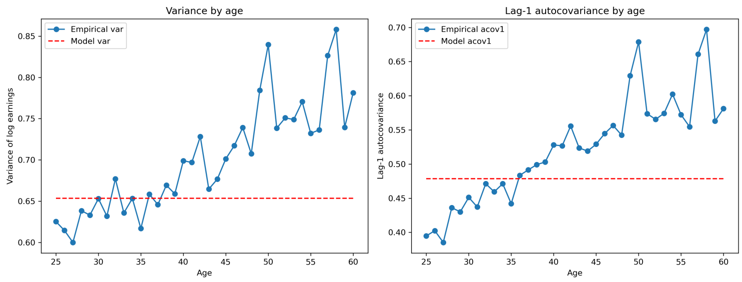
More broadly, these results demonstrate that the sectoral composition of local labour markets serves as a critical determinant of household financial behaviour. As technological change unevenly shifts labour demand across industries, the geographic distribution of income risk, and the resulting pattern of precautionary savings, will evolve in ways that aggregate uncertainty metrics fail to capture.

7 APPENDIX

7.1 Additional model-fit figures

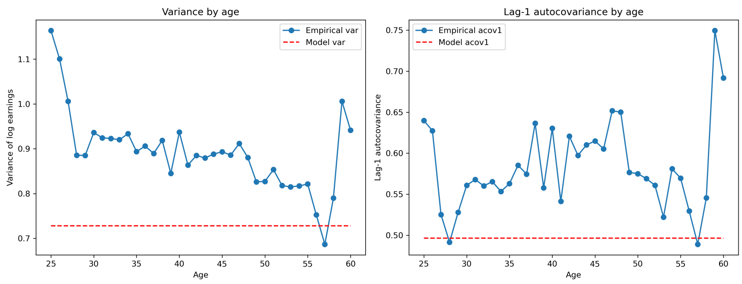
Figures 4–6 report the empirical and model-implied moment profiles for the three remaining sectors estimated in Table 1.

Figure 4: Variance and Lag-1 Autocovariance of Log Earnings, the Trade Sector



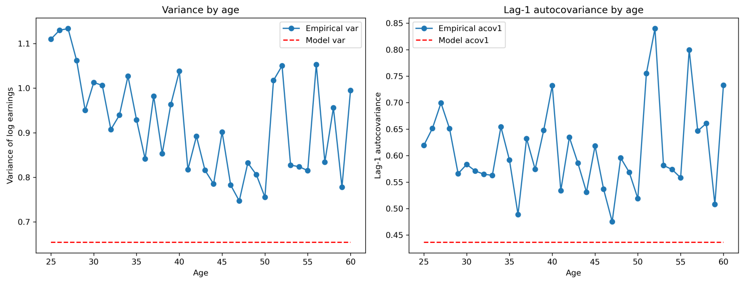
Notes: The left panel plots the cross-sectional variance of log earnings by age; the right panel plots the lag-1 (one-month) autocovariance by age. Solid blue lines with markers are empirical moments computed from the SIPA sample; red dashed lines are the model-implied moments evaluated at the two-step GMM estimates for Trade reported in Table 1. Ages 25–60. The model-implied profiles are flat because the estimated monthly ρ implies that the persistent component reaches its stationary variance well before age 25; see the discussion in Section 5.1.

Figure 5: Variance and Lag-1 Autocovariance of Log Earnings, the Agriculture Sector



Notes: The left panel plots the cross-sectional variance of log earnings by age; the right panel plots the lag-1 (one-month) autocovariance by age. Solid blue lines with markers are empirical moments computed from the SIPA sample; red dashed lines are the model-implied moments evaluated at the two-step GMM estimates for Agriculture reported in Table 1. Ages 25–60. The model-implied profiles are flat because the estimated monthly ρ implies that the persistent component reaches its stationary variance well before age 25; see the discussion in Section 5.1.

Figure 6: Variance and Lag-1 Autocovariance of Log Earnings, the Hotels and Catering Sector



Notes: The left panel plots the cross-sectional variance of log earnings by age; the right panel plots the lag-1 (one-month) autocovariance by age. Solid blue lines with markers are empirical moments computed from the SIPA sample; red dashed lines are the model-implied moments evaluated at the two-step GMM estimates for Hotels and Catering reported in Table 1. Ages 25–60. The model-implied profiles are flat because the estimated monthly ρ implies that the persistent component reaches its stationary variance well before age 25; see the discussion in Section 5.1.

7.2 Multi-Sector Extension

While the baseline model is estimated separately for each sector, the multi-sector extension estimates all K sectors jointly. The key modification replaces the state-dependent scalar innovation variances ($\sigma^2_{z,\text{good}}$, $\sigma^2_{z,\text{bad}}$, π) with a full $K \times K$ covariance matrix for persistent shocks, capturing simultaneously the within-sector innovation variance of each industry and the cross-sector co-movement of persistent earnings risk.

Earnings process. For worker i in sector $k \in \{1, \dots, K\}$ at age h and time t, log earnings follow

$$y^{h,k}_{i,t} = g\Big(x^{h,k}_{i,t},\, Y_t\Big) + u^k_{i,t}, \tag{7}$$

where the idiosyncratic component retains the three-way decomposition of the baseline:

$$u^k_{i,t} = \alpha_i + z^k_{i,t} + \varepsilon^k_{i,t}. \tag{8}$$

As before, $\alpha_i \sim \mathcal{N}(0, \sigma^2_\alpha)$ is a worker-specific permanent component that is common across all sectors, $\varepsilon^k_{i,t} \sim \mathcal{N}(0, \sigma^2_\varepsilon)$ is an i.i.d. transitory shock, and the sector-specific persistent component satisfies the AR(1) process

$$z^k_{i,t} = \rho\, z^k_{i,t-1} + \eta^k_{i,t}, \tag{9}$$

with a common autoregressive coefficient $\rho \in (0, 1)$ imposed across all K sectors.

Cross-sector covariance structure of persistent shocks. Collect the K sector-specific persistent innovations into the vector

$$\boldsymbol{\eta}_{i,t} = \Big(\eta^1_{i,t},\, \eta^2_{i,t},\, \dots,\, \eta^K_{i,t}\Big)^\top.$$

The joint distribution of persistent shocks across sectors is

$$\boldsymbol{\eta}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(\mathbf{0},\, \boldsymbol{\Sigma}_\eta)\,, \tag{10}$$

where $\boldsymbol{\Sigma}_\eta$ is a $K \times K$ symmetric positive definite matrix. The diagonal element $[\boldsymbol{\Sigma}_\eta]_{kk}$ gives the innovation variance of the persistent shock in sector k, while the off-diagonal element $[\boldsymbol{\Sigma}_\eta]_{kk'}$ ($k \neq k'$) captures the contemporaneous co-movement of persistent shocks between sectors k and k'. This structure replaces the two-state parameterization ($\sigma^2_{z,\text{good}}$, $\sigma^2_{z,\text{bad}}$, π) of the baseline with a richer cross-sectional dependence structure.

To ensure positive definiteness and numerical stability during optimization, $\boldsymbol{\Sigma}_\eta$ is parameterized via the Cholesky factorization

$$\boldsymbol{\Sigma}_\eta = \mathbf{L}\mathbf{L}^\top, \tag{11}$$

where $\mathbf{L}$ is a $K \times K$ lower-triangular matrix with strictly positive diagonal entries. The $K(K+1)/2$ free elements of $\mathbf{L}$ are estimated directly: diagonal entries are constrained to $(0, 1]$ and off-diagonal entries are unrestricted in $[-1, 1]$.

Model-implied moments. Let $\boldsymbol{\Sigma}_Z^{(h)}$ be the $K \times K$ matrix whose (k, k') entry is the model-implied cross-sectional covariance of the persistent components $z_{i,t}^k$ and $z_{i,t}^{k'}$ for workers of age h. Initialized at $\boldsymbol{\Sigma}_Z^{(h_{\min})} = \mathbf{0}_{K\times K}$, the matrix is propagated forward through the discrete Lyapunov recursion

$$\boldsymbol{\Sigma}_Z^{(h)} = \rho^2\, \boldsymbol{\Sigma}_Z^{(h-1)} + \boldsymbol{\Sigma}_\eta, \qquad h = h_{\min} + 1,\, \ldots,\, h_{\max}, \tag{12}$$

iterated at annual frequency. Because $\boldsymbol{\Sigma}_\eta$ is age-invariant and the recursion is initialized at the zero matrix, iterating (12) yields the closed-form solution

$$\boldsymbol{\Sigma}_Z^{(h)} = \left(\sum_{j=0}^{h-h_{\min}-1} \rho^{2j} \right) \boldsymbol{\Sigma}_\eta = \frac{1 - \rho^{2(h-h_{\min})}}{1-\rho^2}\, \boldsymbol{\Sigma}_\eta, \qquad |\rho| < 1. \tag{13}$$

Equation (13) shows that $\boldsymbol{\Sigma}_Z^{(h)}$ is proportional to $\boldsymbol{\Sigma}_\eta$ at every age: the cross-sector correlation structure of persistent shocks is preserved across the life cycle, while the overall scale of earnings dispersion grows with worker age.

Given $\boldsymbol{\Sigma}_Z^{(h)}$, the model-implied moments are as follows. The *within-sector variance* for sector k at age h is

$$\mathrm{Var}\Big(y_{i,t}^{h,k}\Big) = \sigma_\alpha^2 + \Big[\boldsymbol{\Sigma}_Z^{(h)}\Big]_{kk} + \sigma_\varepsilon^2, \tag{14}$$

and the *cross-sector covariance* between sectors $k \neq k'$ at age h is

$$\mathrm{Cov}\Big(y_{i,t}^{h,k},\, y_{i,t}^{h,k'}\Big) = \sigma_\alpha^2 + \Big[\boldsymbol{\Sigma}_Z^{(h)}\Big]_{kk'}, \tag{15}$$

where $[\mathbf{A}]_{kk'}$ denotes the (k, k') entry of matrix $\mathbf{A}$. The fixed-effect variance σ_α^2 enters both expressions because workers carry their permanent earnings component α_i across sectoral affiliations, contributing to both within-sector dispersion and cross-sector co-movement. Transitory shocks $\varepsilon_{i,t}^k$ are sector- and period-specific and therefore do not contribute to the cross-sector covariance in (15).

Empirical moments. Two sets of sample moments are formed for each age $h \in \{h_{\min}, \ldots, h_{\max}\}$.

1. *Within-sector variance.* For each sector k, the cross-sectional sample variance of individual log earnings within the sector–age cell:

$$\hat{V}^k(h) = \widehat{\mathrm{Var}}\left(y_{i,t}^{h,k}\right).$$

2. *Cross-sector covariance.* For each ordered pair (k, k') with $k' < k$, the sample covariance of sector-mean log earnings computed across time periods within each age group:

$$\hat{C}^{kk'}(h) = \widehat{\mathrm{Cov}}\left(\bar{y}_t^k(h),\ \bar{y}_t^{k'}(h)\right),$$

   where $\bar{y}_t^k(h) = \left|\mathcal{S}_{t,h}^k\right|^{-1} \sum_{i \in \mathcal{S}_{t,h}^k} y_{i,t}^{h,k}$ is the mean log earnings of age-h workers in sector k at time t, and $\mathcal{S}_{t,h}^k$ denotes the set of workers of age h employed in sector k at time t.

Stacking the K within-sector variances and the $K(K-1)/2$ unique cross-sector covariances over all age groups, the total number of moment conditions is

$$M = \frac{K(K+1)}{2} \times \left(h_{\max} - h_{\min} + 1\right). \tag{16}$$

For $K = 15$ sectors and an age range from 25 to 60 years, this yields $M = 120 \times 36 = 4{,}320$ moment conditions. These replace the variance and lag-1 autocovariance age profiles used in the single-sector baseline and leverage the cross-sectoral dimension of the data to identify every element of $\mathbf{\Sigma}_\eta$.

Parameter vector. The full parameter vector of the multi-sector model is

$$\boldsymbol{\theta} = \left(\rho,\ \sigma_\alpha,\ \sigma_\varepsilon,\ \mathrm{vech}(\mathbf{L})\right)^\top, \tag{17}$$

where $\mathrm{vech}(\mathbf{L})$ collects the $K(K+1)/2$ lower-triangular elements of $\mathbf{L}$ in column-major order. The total number of free parameters is $3 + K(K+1)/2$. For $K = 15$ sectors this amounts to $3 + 120 = 123$ parameters, compared with the six parameters of the single-sector baseline. Identification requires $M > 3 + K(K+1)/2$, a condition easily satisfied for any age range of practical length.

Estimation. The two-step GMM procedure described in Section 3.1 applies without modification. The first step uses the identity weighting matrix $\mathbf{W}_0 = \mathbf{I}_M$. For the second step, the moment covariance matrix is estimated by the regularized outer product of first-step residuals,

$$\hat{\mathbf{S}} = \mathbf{r}\,\mathbf{r}^\top + \lambda\,\mathbf{I}_M, \qquad \mathbf{r} = \hat{\mathbf{m}}_{\text{data}} - \mathbf{m}_{\text{model}}\Big(\hat{\boldsymbol{\theta}}^{(1)}\Big), \tag{18}$$

with regularization parameter $\lambda = 10^{-4}$, and the optimal weighting matrix is $\mathbf{W}^* = \hat{\mathbf{S}}^{-1}$. Regularization prevents near-singularity that arises when M is large relative to the time dimension of the panel. The asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}^{(2)}$ is computed via the sandwich formula of Section 3.1, using the multi-sector Jacobian $\mathbf{G}(\boldsymbol{\theta}) = \partial\,\mathbf{m}_{\text{model}}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^\top$ evaluated at $\hat{\boldsymbol{\theta}}^{(2)}$.

Table 5: Baseline Single-Sector Model vs. Multi-Sector Extension

Feature	**Baseline (per sector)**	**Multi-sector (K sectors jointly)**
Persistent shock variance	Scalar, state-dependent: $\sigma^2_{z,s}$, $s \in \{\text{good}, \text{bad}\}$	$K \times K$ matrix $\boldsymbol{\Sigma}_\eta = \mathbf{L}\mathbf{L}^\top$
Business-cycle states	Yes: $\sigma^2_{z,\text{good}}$, $\sigma^2_{z,\text{bad}}$, π	Removed; absorbed into cross-sector structure
Lyapunov recursion	Scalar: $\sigma^2_z \leftarrow \rho^2\sigma^2_z + \sigma^2_{\eta,s}$	Matrix: $\boldsymbol{\Sigma}_Z^{(h)} \leftarrow \rho^2\boldsymbol{\Sigma}_Z^{(h-1)} + \boldsymbol{\Sigma}_\eta$
Moments matched per age	2 ($\hat{V}(h)$ and $\widehat{\text{ACov}_1}(h)$)	$K(K+1)/2$ (variances + cross-covariances)
Total moment conditions	$2\,(h_{\max} - h_{\min} + 1)$	$\dfrac{K(K+1)}{2}\,(h_{\max} - h_{\min} + 1)$
Free parameters	6 $(\rho, \sigma_\alpha, \sigma_\varepsilon, \sigma_{z,\text{good}}, \sigma_{z,\text{bad}}, \pi)$	$3 + K(K+1)/2$ $(\rho, \sigma_\alpha, \sigma_\varepsilon, \text{vech}(\mathbf{L}))$
For $K = 15$ sectors, $h \in [25, 60]$:		
Parameters	6	$3 + 120 = 123$
Moment conditions	$2 \times 36 = 72$	$120 \times 36 = 4{,}320$
Weighting matrix $\mathbf{W}^*$	$\hat{\mathbf{S}}^{-1}$, unregularized	$\hat{\mathbf{S}}^{-1}$ regularized ($\lambda = 10^{-4}$)

8 Quasi-Bayesian GMM Estimation

To quantify the structural persistence, variance components, and cross-sector spillover variances, we estimate the model using a Quasi-Bayesian Generalized Method of Moments (Laplace Moment Likelihood) framework. This approach combines the structural dynamic moment conditions of the model with Markov Chain Monte Carlo (MCMC) sampling, yielding full posterior distributions for the structural parameter vector $\boldsymbol{\theta}$.

8.1 Parameter Vector and Structural Moments

Let $\mathcal{K} = \{0, 1, \ldots, K-1\}$ denote the set of $K = 15$ sectors. The parameter vector of interest is defined as:

$$\boldsymbol{\theta} = \Big(\rho, \sigma_\alpha, \sigma_\varepsilon, \mathrm{vech}(\mathbf{L})\Big)^\top \tag{19}$$

where $\rho \in (0,1)$ measures the persistence of the AR(1) shock, σ_α is the standard deviation of individual fixed effects, σ_ε is the transitory error standard deviation, and $\mathbf{L}$ is a lower triangular matrix derived from the Cholesky decomposition of the structural shock covariance matrix, $\boldsymbol{\Sigma}_\eta = \mathbf{L}\mathbf{L}^\top$.

The model generates theoretical variance and cross-sector covariance profiles over the worker's life cycle $h \in \{h_{\min}, \ldots, h_{\max}\}$. The dynamic state covariance matrix $\boldsymbol{\Sigma}_Z(h)$ evolves according to the generalized matrix Lyapunov equation:

$$\boldsymbol{\Sigma}_Z(h) = \rho^2 \boldsymbol{\Sigma}_Z(h-1) + \boldsymbol{\Sigma}_\eta, \quad \text{with } \boldsymbol{\Sigma}_Z(0) = \mathbf{0}_{K\times K} \tag{20}$$

For each age h, the theoretical sector-specific variances and cross-sector covariances are given by:

$$\mathrm{Var}(y_{i,s,h} \mid \boldsymbol{\theta}) = \sigma_\alpha^2 + [\boldsymbol{\Sigma}_Z(h)]_{s,s} + \sigma_\varepsilon^2, \quad \forall s \in \mathcal{K} \tag{21}$$

$$\mathrm{Cov}(y_{i,s_1,h}, y_{i,s_2,h} \mid \boldsymbol{\theta}) = \sigma_\alpha^2 + [\boldsymbol{\Sigma}_Z(h)]_{s_1,s_2}, \quad \forall s_1 > s_2 \tag{22}$$

Stacking these moments across ages yields the theoretical moment vector $\mathbf{m}(\boldsymbol{\theta})$ of size $M = K(K+1)/2 \times (h_{\max} - h_{\min} + 1)$.

8.2 Quasi-Likelihood and Posterior Sampling

Let $\hat{\mathbf{m}}$ denote the vector of empirical moments computed from the panel sample of size N. Following Chernozhukov and Hong (2003), we form a Laplace quasi-likelihood based on the GMM quadratic criterion:

$$\mathcal{L}_n(\hat{\mathbf{m}} \mid \boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}\left[\hat{\mathbf{m}} - \mathbf{m}(\boldsymbol{\theta})\right]^\top \mathbf{W} \left[\hat{\mathbf{m}} - \mathbf{m}(\boldsymbol{\theta})\right]\right) \tag{23}$$

where $\mathbf{W}$ is the GMM weighting matrix.

The posterior density $p(\boldsymbol{\theta} \mid \hat{\mathbf{m}})$ is proportional to the product of the quasi-likelihood and regularizing priors $p(\boldsymbol{\theta})$:

$$p(\boldsymbol{\theta} \mid \hat{\mathbf{m}}) \propto p(\boldsymbol{\theta}) \cdot \exp\left(-\frac{1}{2}Q_N(\boldsymbol{\theta})\right) \tag{24}$$

where $Q_N(\boldsymbol{\theta}) = [\hat{\mathbf{m}} - \mathbf{m}(\boldsymbol{\theta})]^\top \mathbf{W}[\hat{\mathbf{m}} - \mathbf{m}(\boldsymbol{\theta})]$.

8.3 Priors and MCMC Implementation

We impose parameter bounds through truncated normal priors $p(\boldsymbol{\theta})$:

- Persistence: $\rho \sim \mathcal{N}(0.50, 0.25^2) \cdot \mathbb{I}(0.01 < \rho < 0.99)$
- Variance components: $\sigma_\alpha \sim \mathcal{N}^+(0.30, 0.50^2)$ and $\sigma_\varepsilon \sim \mathcal{N}^+(0.30, 0.50^2)$
- Cholesky elements: $L_{s,s} \sim \mathcal{N}^+(0.10, 0.50^2)$ for diagonal elements to guarantee positive definiteness of $\boldsymbol{\Sigma}_\eta$, and $L_{s_1,s_2} \sim \mathcal{N}(0.00, 0.50^2)$ for off-diagonal terms.

Posterior sampling is performed using an Affine Invariant Ensemble MCMC sampler (`emcee`; Foreman-Mackey et al., 2013) initialized with $R = 300$ walkers over $S = 1{,}500$ steps per walker, discarding the first 500 steps as burn-in.