Global Income Dynamics Project: Mexico

M. Calderón, A. Cebreros, L. Fernández, J. A. Inguanzo,
D. Jaume, and **D. Puggioni***

November 20 2020

*The views and conclusions presented in this work are exclusively of the authors and do not necessarily reflect those of Banco de México or its Board of Governors.

The Mexican administrative data

- The analysis is based on administrative records from the Instituto Mexicano de Seguridad Social (IMSS):
  - Universe of **formal** private sector workers
    - No information on informal workers
  - Between 13 million and 20 million monthly observations
  - Period from 2005 to 2019
  - Worker's daily taxable income ("base contribution salary")
    - may include various forms of compensation in addition to wages, not necessarily corresponding to total labor income
  - Bottom and top coding
    - Bottom code: minimum wage (1.3% of observations)
    - Top code: 25 minimum wages before February 2017 and 25 UMA (unit of measurement and update) afterwards (1.7% of the observations)

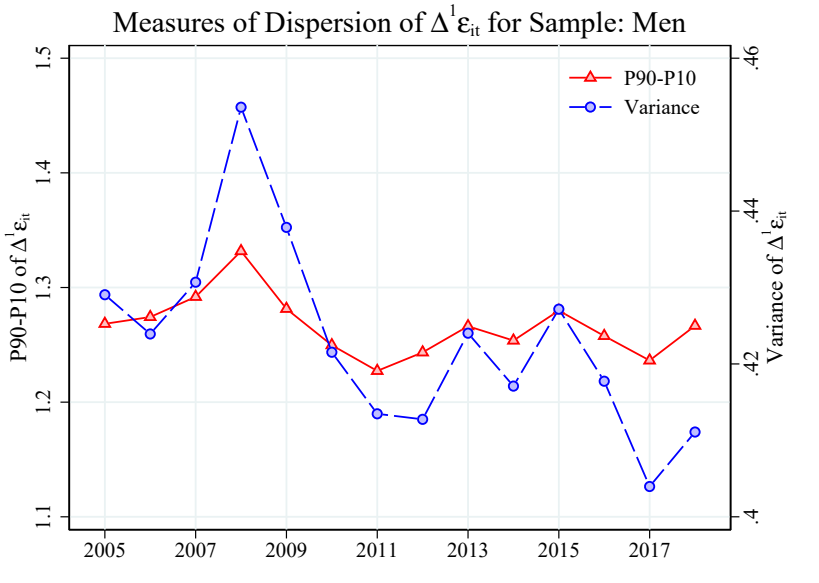
The Mexican administrative data

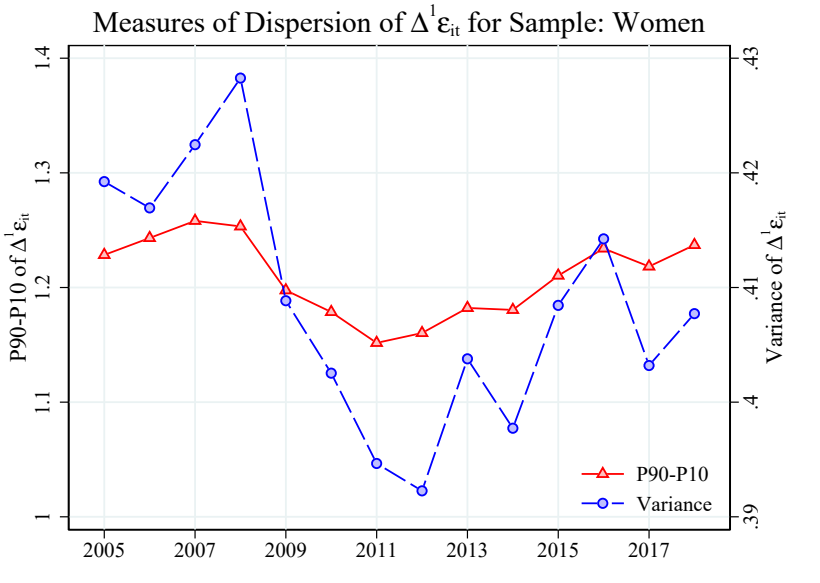
- Information available
  - About the **worker**:
    - SSN, year of birth, year of first enrollment
    - gender
    - type of employment (permanent vs. temporary)
    - daily taxable wage
  - About the **employer**:
    - employer ID, firm tax ID (starting from 11/2018)
    - sector of economic activity
    - location (county) of the employer
- The **meaningful attachment** to the labor force is defined as 45 days of minimum wage

One-year income shocks: main findings

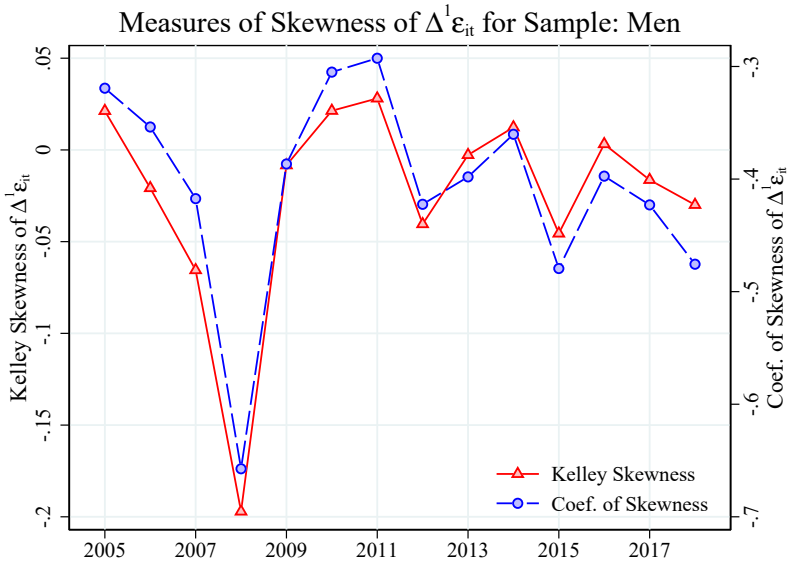
- The analysis of one-year changes in (residualized) income reveals that the distribution of income shocks is:
  - Characterized by decreasing variance but fairly stable P90-P10
  - Negatively skewed
  - Displays excess kurtosis
  - Pareto distribution at both tails, with left tail declining more slowly than the right
  - More volatile and with a greater deviation from normality for younger and lower (permanent) income workers
  - Asymmetry is similar for men and women; kurtosis is higher for women than for men; dispersion seems to be marginally lower for women

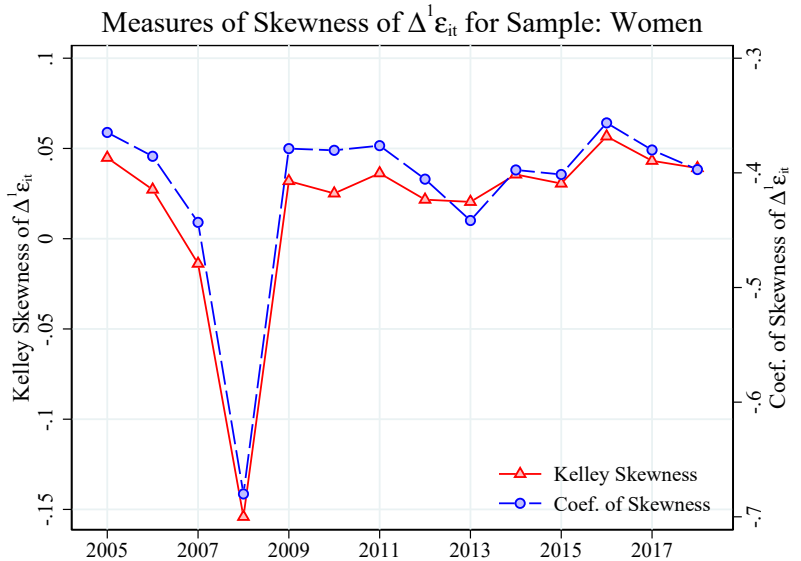
One-year income shocks: dispersion



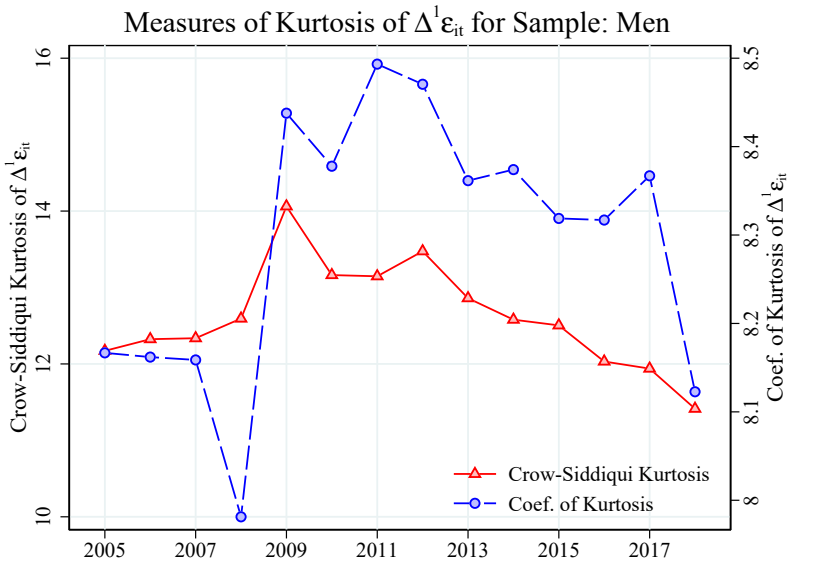


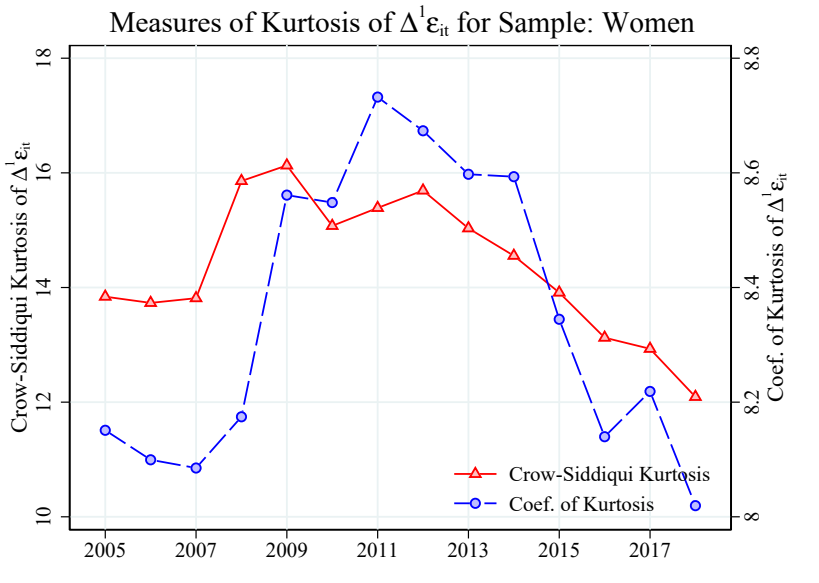
One-year income shocks: skewness



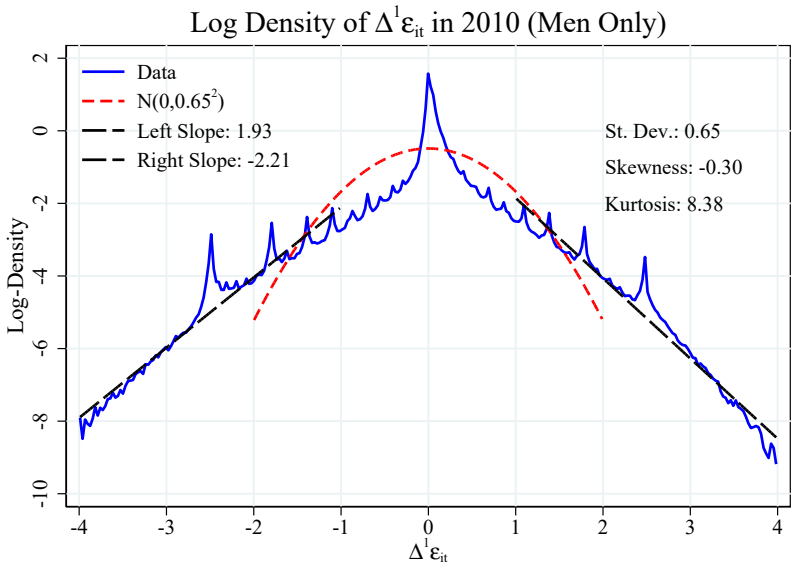


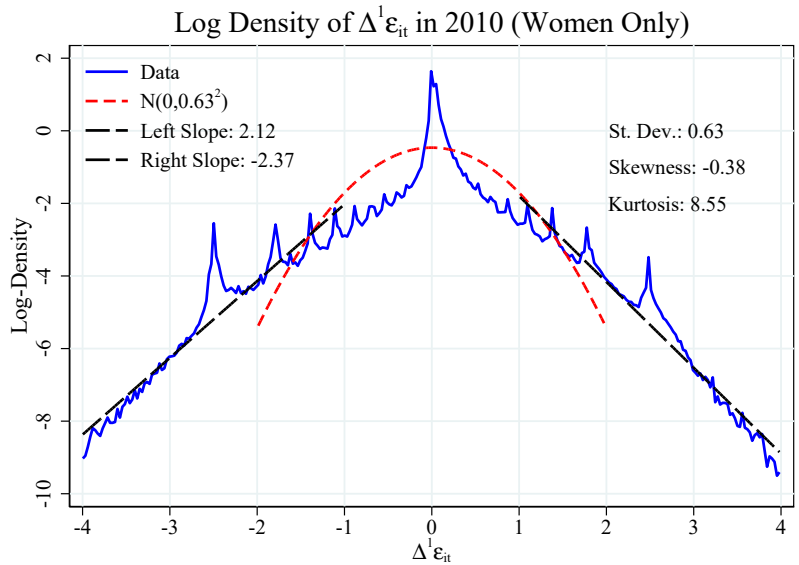
One-year income shocks: kurtosis



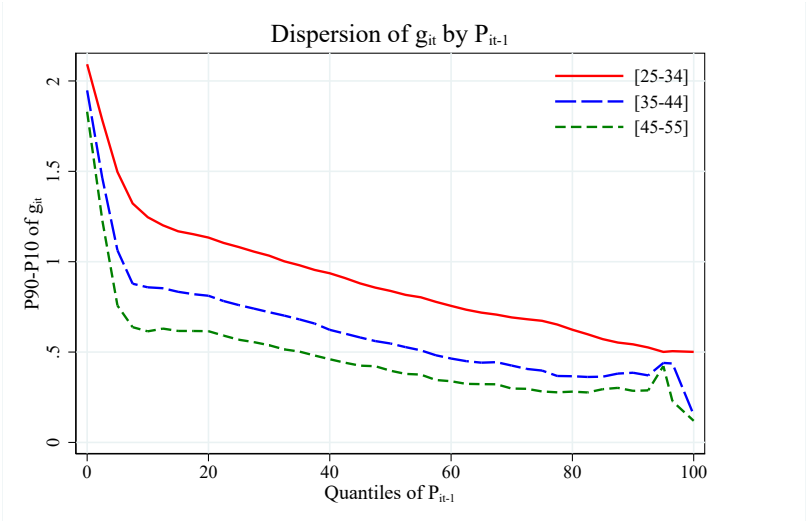


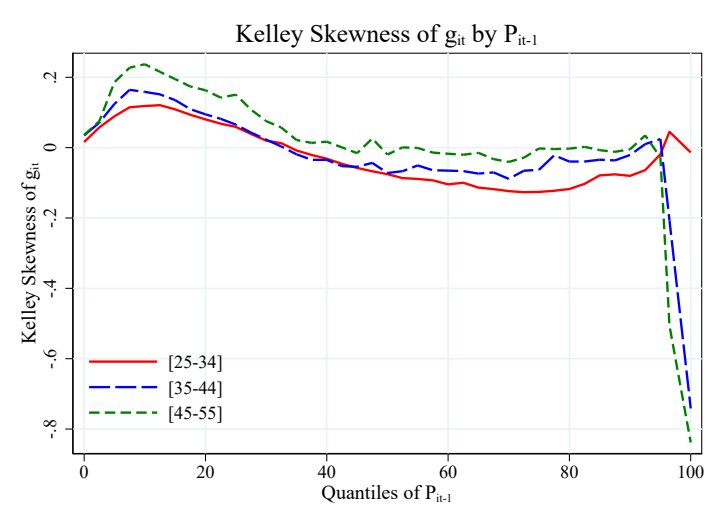
One-year income shocks: log density

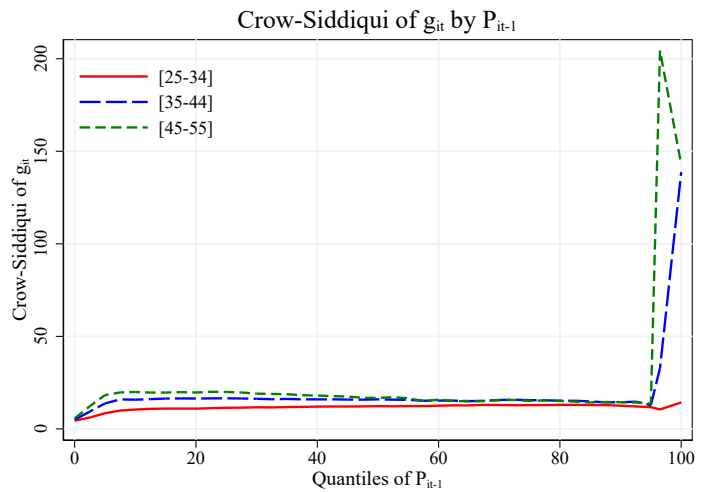




One-year income shocks: |on permanent income



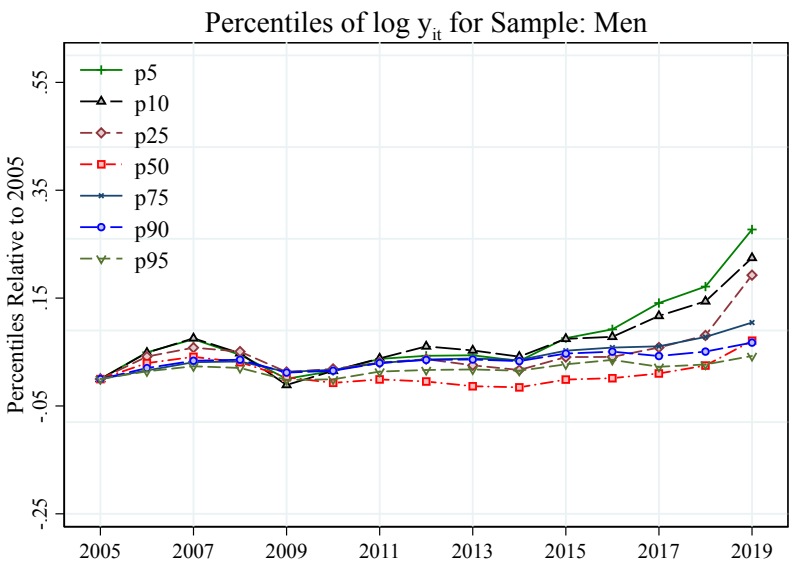


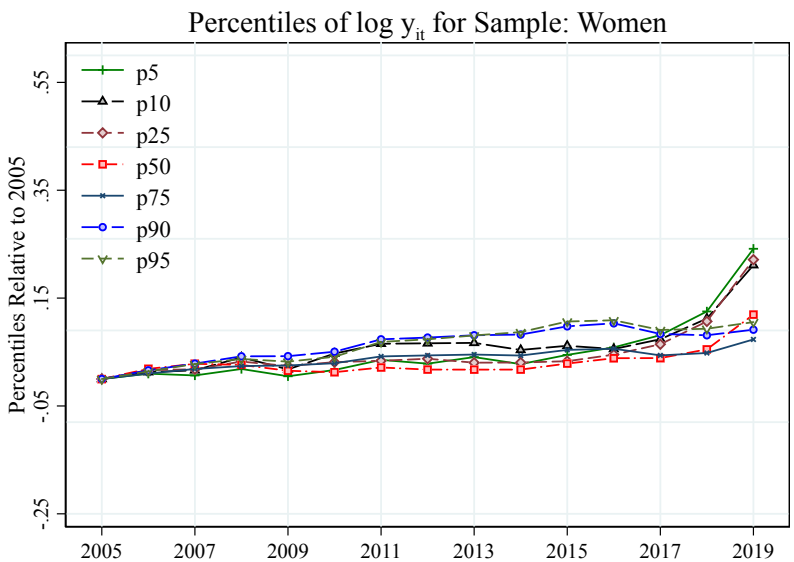


Inequality: log earnings

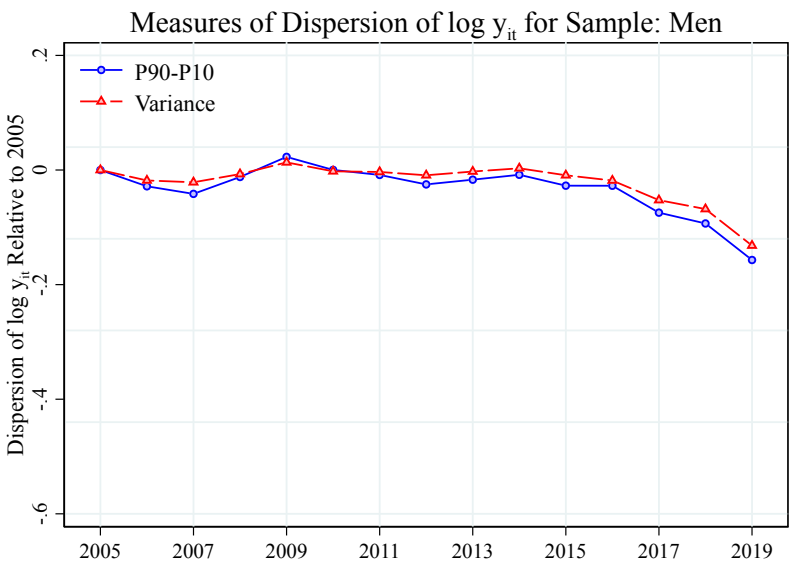
- Inequality in log income has decreased, for both men and women
  - For variance and P90-P10, most of this decreases occurs from 2014 onwards
  - For P90-P50, there is an increase until 2016, followed by a steady decrease
  - For P50-P10, there is a strong downward trend throughout the sample period
  - The reduction in log income inequality is mainly due to the strong growth in income of the lower percentiles and the relative stability of the income of the highest percentiles
    - Income growth of the lowest deciles can be explained by the dynamics of the minimum wage in recent years

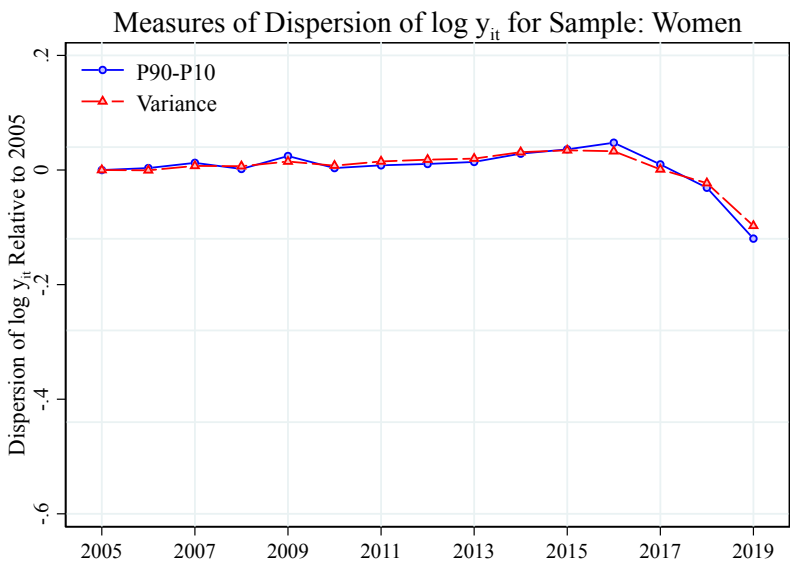
Log earnings: percentiles



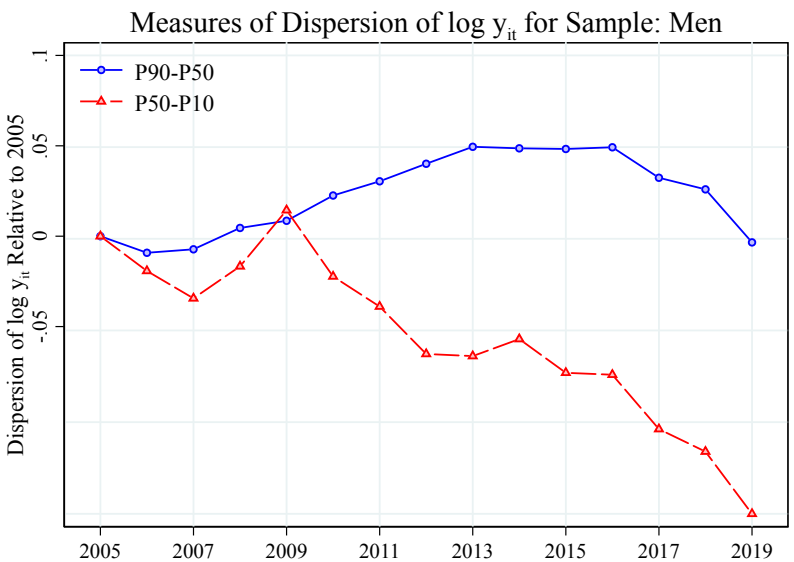


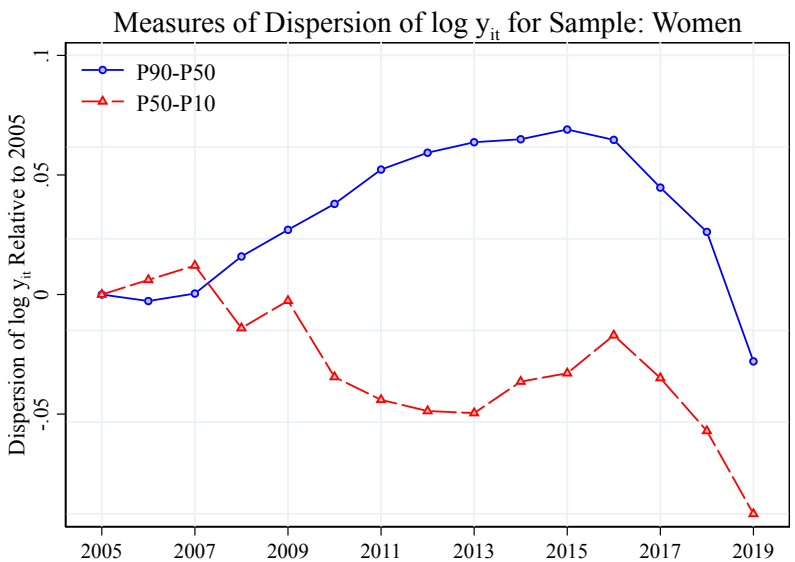
Log earnings: measures of dispersions





Log earnings: dispersion at the top and bottom

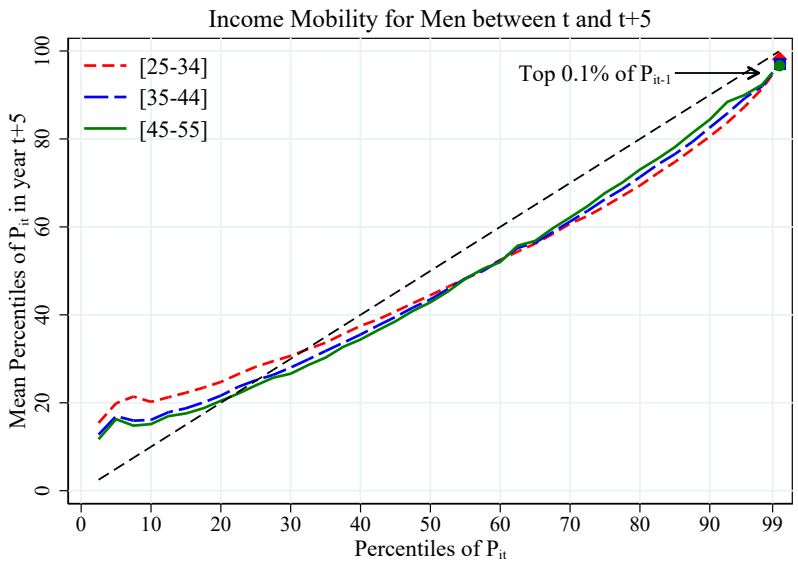


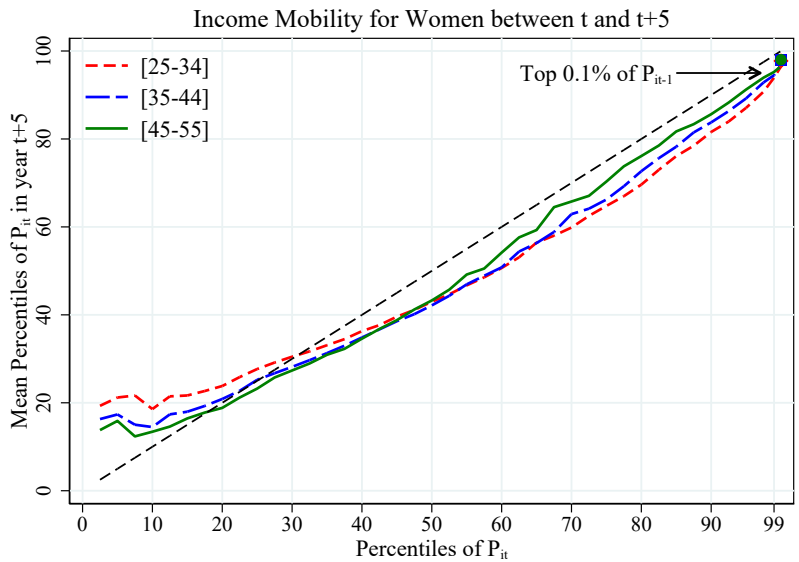


Short- and long-term mobility: main findings

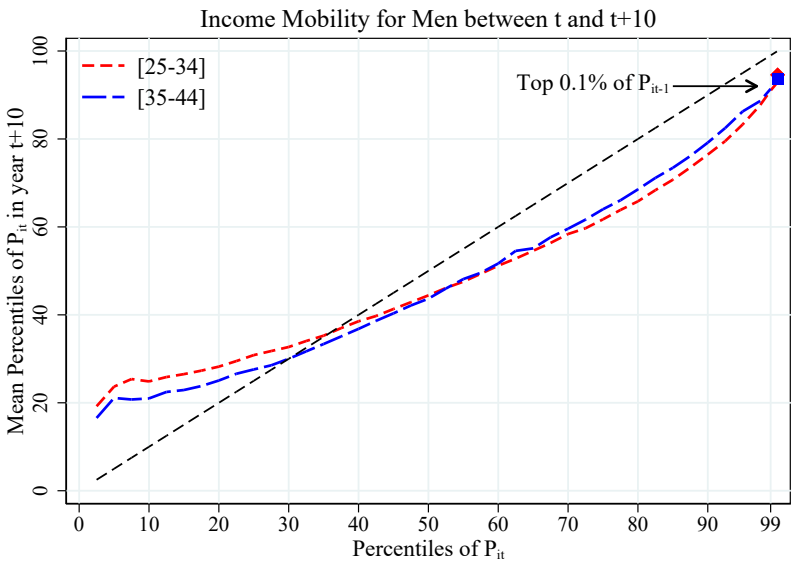
- Upward income mobility at the bottom of the distribution; downward income mobility at the top
- For both short-term and long-term income mobility
  - Upward income mobility up to P20/P30 (depending on age group)
  - Upward income mobility is higher for younger workers
  - From P30 onwards there is downward income mobility
  - Downward income mobility is more likely for younger workers
  - For top 0.1% of permanent income, little to no income mobility in the short-term, but downward income mobility in the long-term.
  - Patterns are qualitatively similar for men and women

Short-term income mobility





Long-term income mobility



Administrative vs. survey data

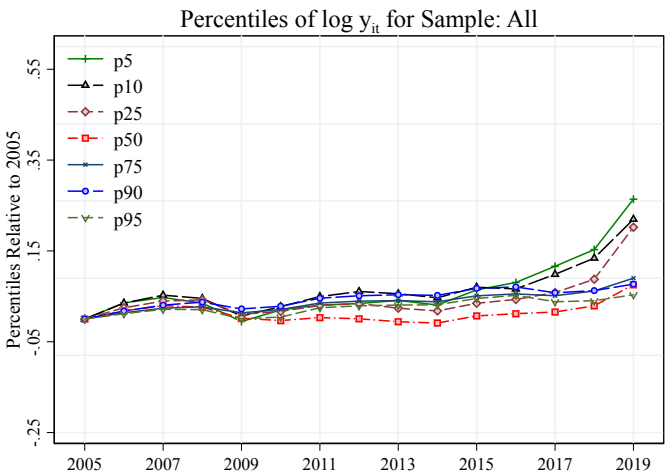
- Administrative data (IMSS) only covers a fraction of the entire labor market
- The informal sector in Mexico absorbs, on average, 60% of the labor force
  - We contrast the statistics of Part A with the same set of statistics constructed with information from HH survey
- Mexico's household survey (ENOE) is a quarterly survey that collects data on the employment situation of Mexicans in rural and urban areas
  - Rotating panel structure, covering the period between the first quarter of 2005 and the first quarter of 2020
  - One-fifth of the sample replaced every quarter
  - Representative at the national and state level (and for selected cities, one for each state)
  - Visits approximately 120 thousand homes in each quarter

Data comparison: main findings

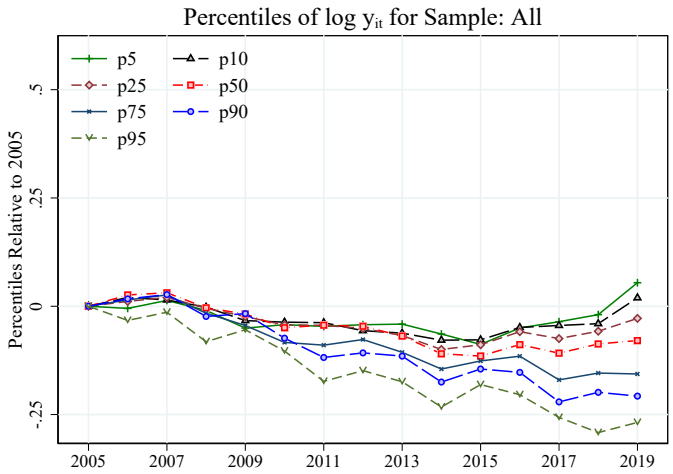
- IMSS vs. ENOE data comparison delivers mixed results:
  - **Percentiles of log income**
    - IMSS: stable throughout sample period, with the exception of lowest percentiles showing an upward trend from 2016
    - ENOE: downward trend throughout sample period; lowest percentiles start trending upward in 2016
  - **Dispersion of log income**
    - IMSS: stable until 2016, trends downward afterwards
    - ENOE: downward trend throughout sample
  - **Percentiles of one-year income changes**
    - IMSS: stable throughout, with the exception of a noticeable decrease during the period of the financial crisis
    - ENOE: stable throughout
  - **Dispersion of one-year income changes**
    - IMSS: sharp increase, followed by sharp decrease during the financial crisis; from 2011 onwards fairly stable variance but p90-p10 trended upwards
    - ENOE: fairly stable, with slight downward trend

Earnings inequality: IMSS vs. ENOE data

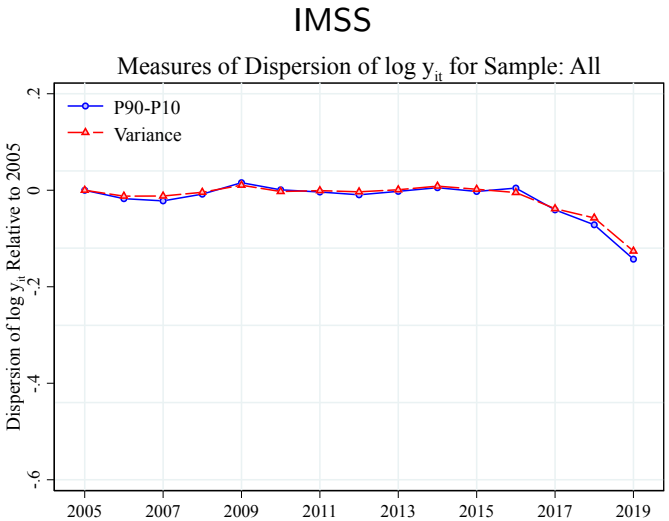
IMSS



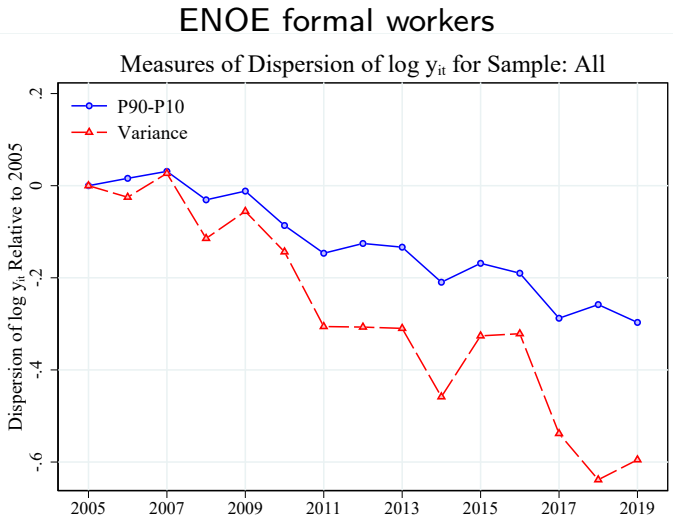
ENOE formal workers



IMSS



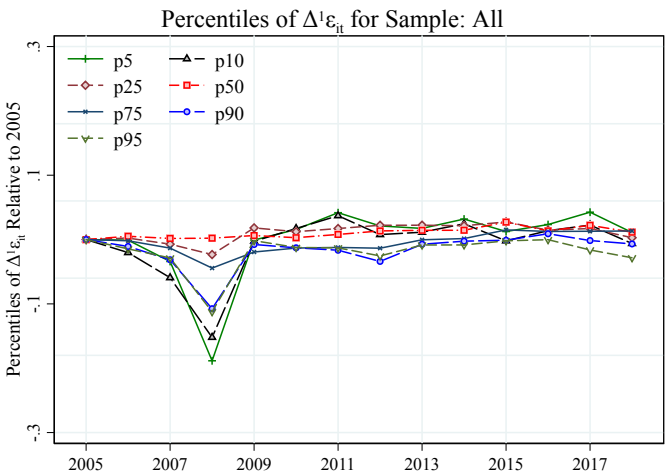
ENOE formal workers



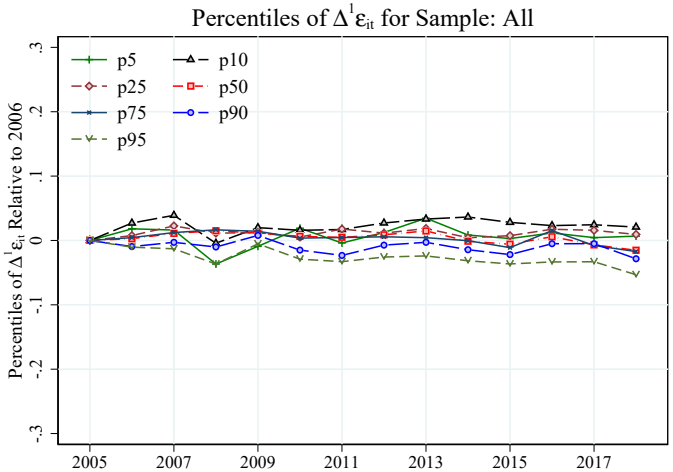
Earnings dynamics: IMSS vs. ENOE data

▸ Part A++

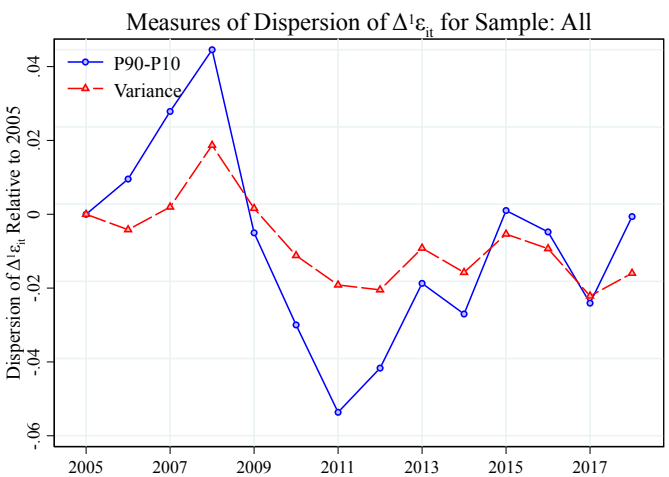
IMSS



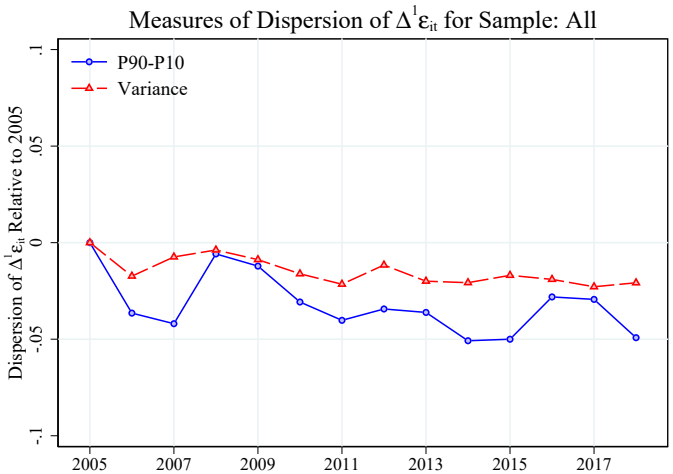
ENOE formal workers



IMSS



ENOE formal workers



Informality in Mexico

- A salient characteristic of developing and emerging market economies is the prevalence of informality
  - In Mexico, between 2005 and 2019, the quarterly rate of informality ranged between 56–60%
  - Levy (2018) shows that informal firms in Mexico represent 90% of all firms and absorb 40% of the economy's capital stock and 55% of employment
  - Informal firms are pervasive in the economy and are not confined to activities deemed "traditional" or "less modern"
- We study two issues that may have important implications for the dynamics of labor income of workers in the Mexican labor market:
  I. The impact of transitions out of and into formal employment on wages earned in the formal sector
  II. The impact of early exposure to informality on labor market outcomes

Part B — main findings

- The dual nature of Mexico's labor market can play an important role in shaping the income dynamics of Mexican workers
  - For workers that **exited the formal labor market**, wages upon re-entry are approximately 15% lower and it can take three or more years to achieve pre-exit wage levels; recovery process faster for women than men
  - For **new cohorts of workers**, entering the labor market during times of higher informality is associated with a higher fraction of the cohort being out of the labor force in the long-term and experiencing negative effect on total earnings
  - Entering the labor market for the **first time into a formal job** can increase future wages by 50 to 80% relative to a worker whose first job was informal, depending on gender and education attainment

Exit and re-entry into the formal sector

- We analyze the wage implications of in-and-out transitions focusing on workers with two formal employment spells
  - Almost one fifth of all formally employed workers have only two spells as formal workers
  - On average, present in the database 3.2 years before leaving, out of formal employment for 2.6 years, come back for another 3.6 years
- Compare the pre-exit wage trajectory with the post re-entry wage trajectory through an event-study approach
  - Baseline approach: 3-year event window before and after exit
  - Balanced panel: only keep workers that we can observe for three consecutive years before leaving and for three consecutive years after re-entry

Exit and re-entry into the formal sector

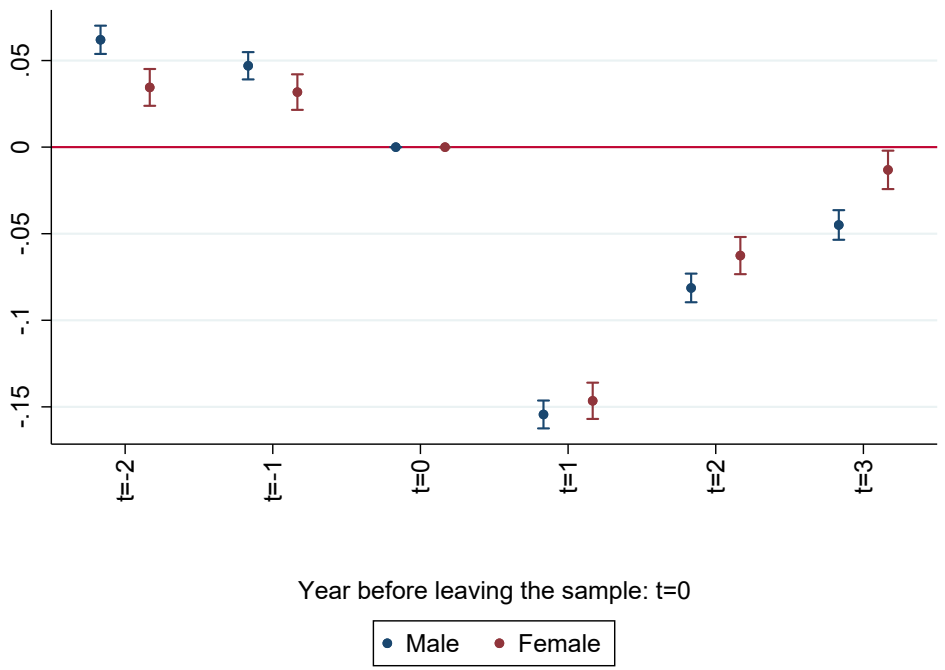
- We compare workers' wage behavior before leaving and after re-entering formal employment by estimating the specification, separately by gender

$$\begin{aligned}\ln(wage)_{it} &= \beta_0 + \sum_{\tau=-2}^{3} \beta_\tau \times \mathbf{1}[event = \tau] + \sum_{k=1}^{5} \beta_k \times \mathbf{1}[duration_i = k] \\ &+ \sum_{\tau=-2}^{3} \sum_{k=1}^{5} \beta_\tau^k \times \mathbf{1}[event = \tau] \times \mathbf{1}[duration_i = k] \\ &+ \mathbf{1}[age_{it} = g] + \alpha_e + \alpha_s + \alpha_t + \varepsilon_{it}\end{aligned}$$

  - where:
    - $\mathbf{1}[event = \tau]$ set of dummy variables for the number of τ years before or after leaving the sample: $\tau = 0$ is the year right before leaving the sample, $\tau = -1$ is two years before leaving the sample, $\tau = 1$ is the first year upon re-entery
    - $\mathbf{1}[duration_i = k]$ set of dummy variables taking the value of 1 depending on the k number of years out of formal employment
    - α_e, α_s, and α_t are fixed effects for sector e, state s, and year t

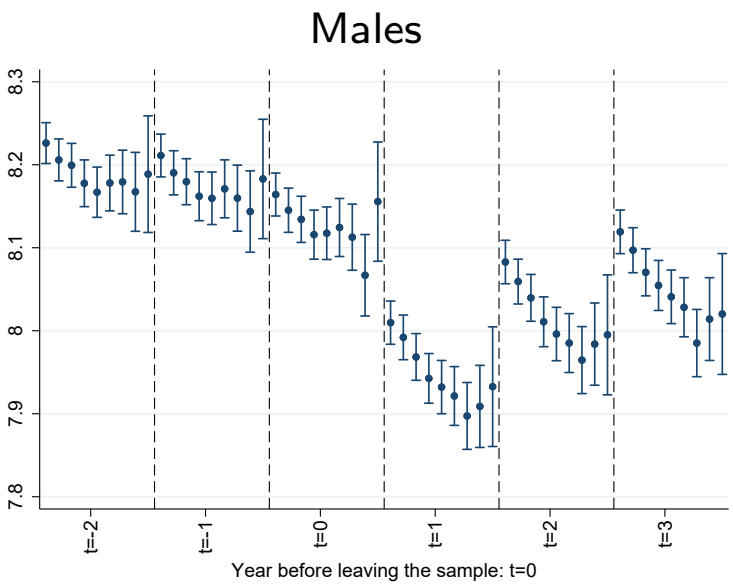
Exit and re-entry into the formal sector

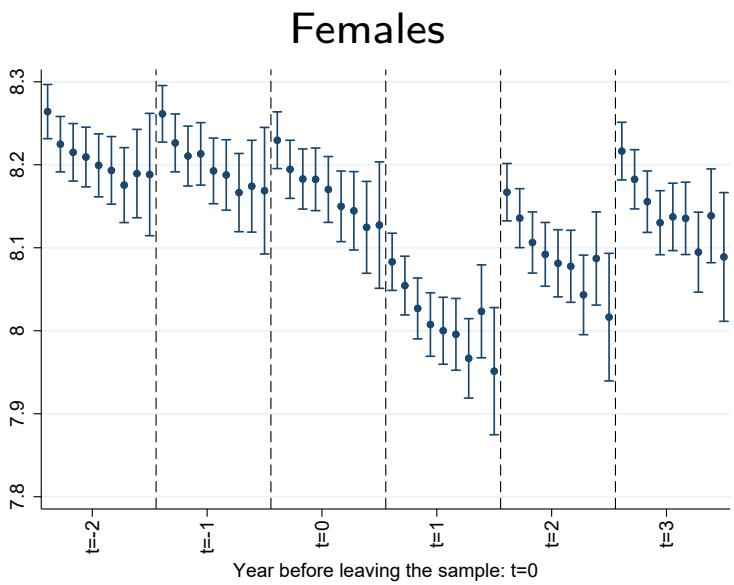
- Wages trajectories of workers who exit and re-enter the formal labor market (log wage differences)
  - Estimated β_τ



Exit and re-entry into the formal sector

- Wages trajectories of workers who exit and re-enter the formal labor market (log wages)
  - Estimated $\beta_0 + \beta_\tau + \beta_k + \beta_\tau^k$





Initial labor market conditions: cohort analysis

- We study the effect of aggregate initial labor market conditions on several labor market outcomes when adults
  - Heterogeneity: separate estimates for workers differentiated by gender and educational attainment
- Initial conditions are unemployment and informality rates
  - calculated at the regional level for workers aged 30 and younger
  - average of the first 3 years after a worker with a given educational level is supposed to have entered the labor force
- Consider cohorts of new workers for which their maximum potential years of experience is at least ten years to study *long-term* outcomes

Initial labor market conditions: cohort analysis

- Following Oreopoulos et al. (2012) we estimate, separately for each demographic group $d \in \{M^h, M^l, W^h, W^l\}$

$$\begin{aligned}\bar{y}^d_{crt} &= \sum_{e=1}^{3} \beta^d_{eu} UR_{cr0} \mathbf{I}[experience^d_{crt} \in e] \\ &+ \sum_{e=1}^{3} \beta^d_{ei} IR_{cr0} \mathbf{I}[experience^d_{crt} \in e] \\ &+ \gamma^d_e + \lambda^d_r + \theta^d_c + \delta^d_t + \mu^d_{crt}\end{aligned}$$

  - where:
    - $\bar{y}^d_{crt}$ is a labor market outcome of demographic group d, belonging to cohort c, in region r, at time t
    - UR_{cr0} and IR_{cr0} are the unemployment rate and the informality rate at the time of entry $t = 0$ for cohort c in region r
    - $experience^d_{crt}$ is years of potential experience of a given demographic group d and e are dummies grouping potential years of experience ($e < 10$, $10 \leq e \leq 14$, $e > 15$ years)
    - γ^d_e, λ^d_r, θ^d_c and δ^d_t are fixed effects by years of potential experience e, region r, cohort c, and year t

Exit and re-entry into the formal sector

Impact of initial labor market conditions on long-term labor force participation

Dependent variable: fraction of cohort not in labor force				
	Low educated		*High educated*	
Independent variables:	Men	Women	Men	Women
Initial informality	0.250***	0.030	0.128	0.254***
	(0.090)	(0.084)	(0.114)	(0.105)
Initial informality×Experience (10–15)	0.025	0.068**	0.038	0.090***
	(0.025)	(0.030)	(0.044)	(0.039)
Initial informality×Experience (>15)	0.059**	0.094***	0.039	0.166***
	(0.028)	(0.036)	(0.039)	(0.038)
Initial unemployment	–0.395	–0.457	–0.305	0.110
	(0.290)	(0.329)	(0.424)	(0.409)
Initial unemployment×Experience (10–15)	–0.245	0.261	0.095	0.311
	(0.236)	(0.255)	(0.385)	(0.299)
Initial unemployment×Experience (>15)	–0.255	–0.184	0.167	0.276
	(0.237)	(0.275)	(0.326)	(0.297)
N. of Observations	5,155	5,123	5,883	5,928

Note: The regression includes potential experience, region, cohort and year fixed effects. Standard errors (in parenthesis) are clustered at the cohort-region level. The stars indicate significance levels (*p < 0.10, **p < 0.05, ***p < 0.01).

Exit and re-entry into the formal sector

Impact of initial labor market conditions on long-term labor earnings

Dependent variable: log total earnings				
	Low educated		*High educated*	
Independent variables:	Men	Women	Men	Women
Initial informality	–0.453	–0.259	–0.240	–1.415***
	(0.336)	(0.473)	(0.315)	(0.396)
Initial informality×Experience (10–15)	–0.031	–0.413**	–0.001	0.039
	(0.111)	(0.163)	(0.116)	(0.114)
Initial informality×Experience (>15)	–0.393***	–0.753***	0.031	–0.042
	(0.101)	(0.168)	(0.107)	(0.124)
Initial unemployment	0.349	3.090**	–0.639	–1.998
	(0.860)	(1.548)	(0.983)	(1.331)
Initial unemployment×Experience (10–15)	1.033	–1.688	0.530	0.443
	(0.842)	(1.214)	(0.835)	(1.068)
Initial unemployment×Experience (>15)	1.113*	–0.238	0.874	1.528
	(0.642)	(1.250)	(0.770)	(1.076)
N. of Observations	5,119	5,070	5,630	5,674

Note: Labor earnings are computed with zeros to avoid conditioning on a potential outcome (labor force participation). Omitting zeros could result in selection bias since the composition of the labor force may change due to the treatment. The variable is constructed as the natural logarithm of the mean total income for each cell of cohort c, demographic group d, region r, and time t. The regression includes potential experience, region, cohort and year fixed effects. Standard errors (in parenthesis) are clustered at the cohort-region level. The stars indicate significance levels (*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$).

Special module of the household survey

- Special module of the ENOE, the Labor Trajectory Module (MOTRAL)
  - Conducted on a subsample from ENOE Q2 of 2012 and 2015
  - Provides information for labor market trajectories over the previous 5 years
  - Information collected at the level of an individual's job
    - Month and year of job's beginning and end
    - Position within the job (employer, employee, self-employed)
    - Monthly wage
    - Economic sector of the job
    - Social security institution (if any)
  - The survey asks respondents about their first job (even if it took place before the years covered by the survey)

Initial labor market conditions: individual analysis

- Use MOTRAL to quantify how entering the labor market for the first time in a formal/informal job affects future wages
- Construct an annual panel from MOTRAL and estimate, separately for each demographic group $d \in \{M^h, M^l, W^h, W^l\}$

$$\begin{aligned}\ln(y_{it}) &= \beta_0 + \beta_1 age_{it} + \beta_2 age_{it}^2 + \beta_3 formal_{i0} \\ &+ \beta_4 (potential\ experience_{it} \times formal_{i0}) \\ &+ \phi_t + \gamma_s + \lambda_{it} + u_{it}\end{aligned}$$

  - where:
    - $\ln(y_{it})$ is the average monthly income
    - age_{it} is the age of individual i in year t
    - $potential\ experience_{it}$ are years of potential experience of i at t
    - $formal_{i0}$ is a dummy equal to 1 when i's first job was formal (instrumented with measures of local labor market informality)
    - λ_{it} is the Heckman correction term (employment selection)
    - ϕ_t and γ_s are year and sector fixed effects

Initial labor market conditions: individual analysis

Impact of initial informality on future wages

Dependent variable: log wages						
	Low educated			*High educated*		
Independent variables:	Men	Women	All	Men	Women	All
Potential experience×Formal	0.044***	0.002	0.042***	0.015	0.077***	0.047***
	(0.015)	(0.014)	(0.011)	(0.013)	(0.016)	(0.009)
Formal	0.026	0.774***	0.186	0.497***	0.792***	0.413***
	(0.139)	(0.216)	(0.113)	(0.126)	(0.236)	(0.110)
Age	0.033***	–0.032	–0.013	0.053**	–0.004	0.041**
	(0.012)	(0.022)	(0.011)	(0.021)	(0.028)	(0.016)
Age^2	–0.001***	0.000	–0.000	–0.001*	–0.000	–0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
λ	–0.116***	–0.140***	–0.292***	–0.129***	–0.301***	–0.345***
	(0.040)	(0.038)	(0.028)	(0.035)	(0.054)	(0.031)
Constant	8.058***	8.654***	8.747***	7.295***	9.332***	8.306***
	(0.187)	(0.503)	(0.189)	(0.334)	(0.489)	(0.289)
N. of Observations	2,545	2,093	4,638	3,937	3,745	7,682

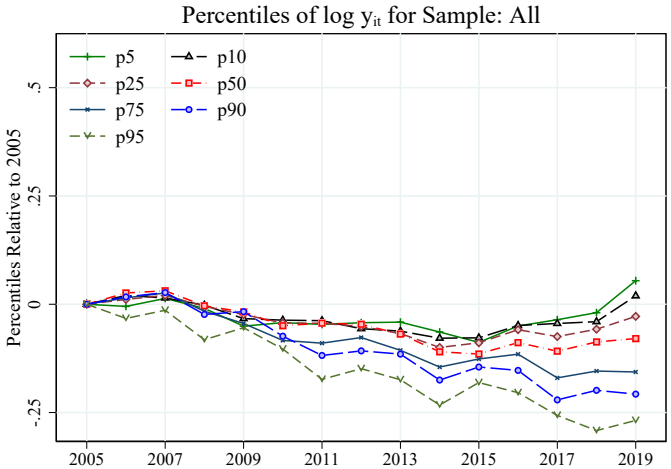
Note: The regression includes sector and year fixed effects. Robust standard errors (in parenthesis). The stars indicate significance levels (*p < 0.10, **p < 0.05, ***p < 0.01).

Summing up

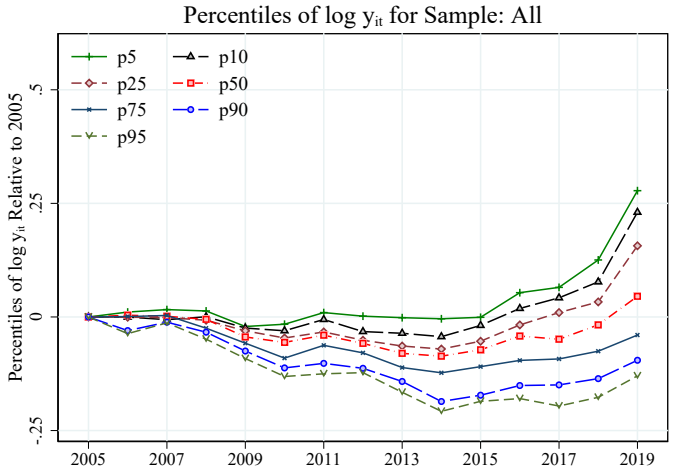
- **Part A**
  - Income shocks are:
    - less dispersed overtime
    - large negative shocks more probable especially for younger and low (permanent) income workers
    - significant mass of workers exposed to large income shocks
    - more tail concentration in negative income shocks
  - Reduction in income inequality
  - Upward income mobility at the bottom and downward mobility at the top with younger workers more affected
- **Part B**
  - Dual nature of labor market in Mexico plays an important role
    - 15% average wage penalty for exiting and re-entering the formal sector
    - higher chance of staying out of the labor force and lower earnings for cohorts entering in period with high informality
    - first job in the formal sector increases future wages between 50 and 80%

Earnings inequality: formal vs. informal (ENOE)

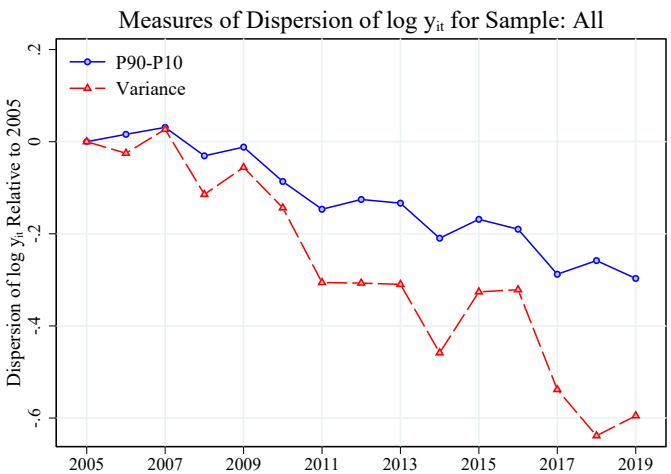
ENOE formal workers



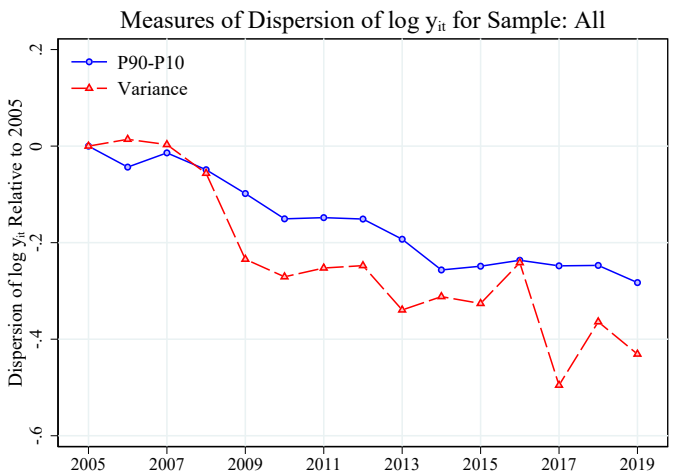
ENOE informal workers



ENOE formal workers

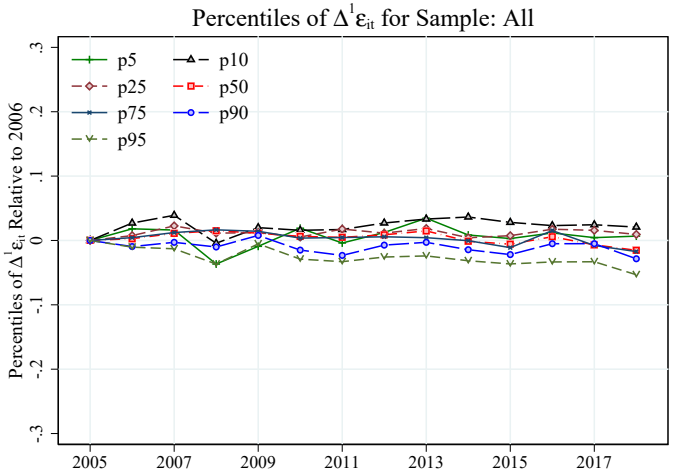


ENOE informal workers

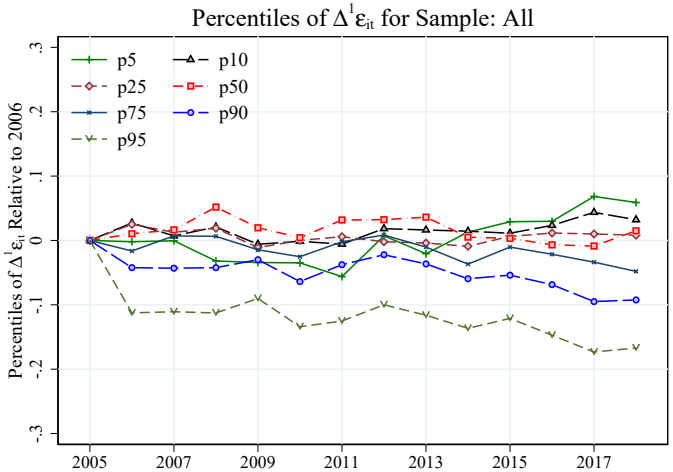


Earnings dynamics: formal vs. informal (ENOE)

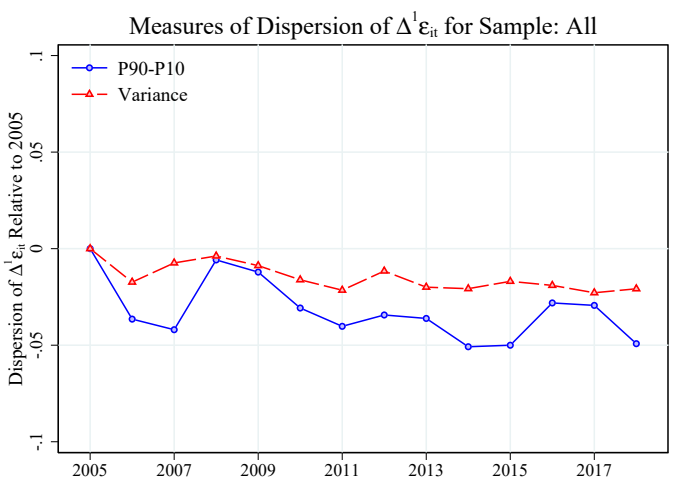
ENOE formal workers



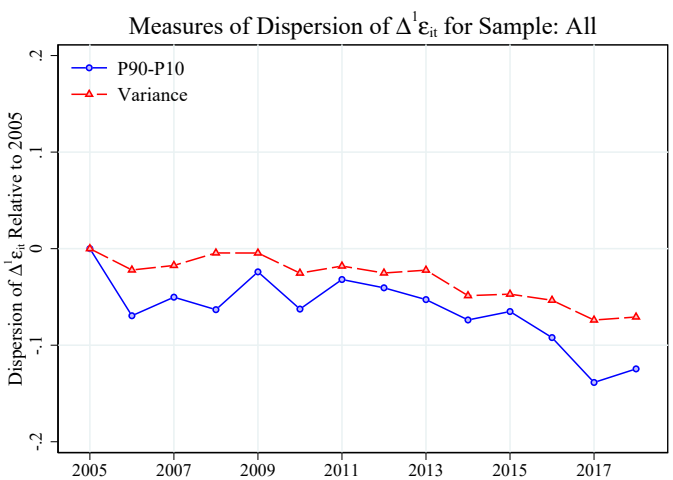
ENOE informal workers



ENOE formal workers



ENOE informal workers



A note of caution regarding survey data

- Calculating the statistics from Part A with household survey data might be problematic for several reasons
  - Measurement error due to the fact that income is self-reported
  - Households may be reporting a broader measure of income than the one captured in the IMSS data
  - Non-response for the income variable: The share of observations with missing incomes in ENOE is, on average, around 25%
    - non-response is not random and is concentrated among formally employed workers and workers with higher levels of educational attainment
  - ENOE is specifically designed to collect information on employment outcomes, not income