

# Are Informal Self-Employment and Informal Employment as Employee Behaviorally Distinct Labor Force States?

Luca Flabbi<sup>1</sup>    Mauricio Tejada<sup>2</sup>

<sup>1</sup>University of North Carolina – Chapel Hill

<sup>2</sup>Universidad Diego Portales and Millenium Nucleus  $LM^2C^2$

*Tejada gratefully acknowledges financial support from FONDECYT grant No 11196296 and ANID – MILENIO – NCS 2022\_045.*

Annual Meeting SECHI 2023

# Motivation

- ▶ In an influential paper published in 1983, Flinn and Heckman asked: “Are Unemployment and out of the Labor Force Behaviorally Distinct Labor Force States?”
  - ▶ **Relevance:** Labor economists had started to study labor market dynamics with richer theoretical and empirical models.
- ▶ The study of labor market dynamics in economies with high informality is experiencing a similar transformation.
  - ▶ **It is necessary to take a stand on:** What are the relevant labor market states? What transitions between labor market states should we focus on?
  - ▶ **Current debate:** Informal employee vs. informal self-employed.

# This Paper

**Question:** Should we differentiate between informal workers who are hired as employees and those who are working as self-employed?

**Why:**

- ▶ Aggregating or differentiating these two labor market states is relevant for estimation and policy.
- ▶ Both approaches have been used by influential papers, without producing a consensus in the literature.

**What do we do:**

- ▶ We conduct non-parametric tests (of equality of empirical distributions).
- ▶ We develop and estimate a search model where we directly impose (and test) the same behavior.

## Our paper in context

- ▶ Literature that estimate (or calibrate) search models aggregating unregistered employees and self-employed or excluding the self-employed.  
[Haanwinckel and Soares (2021); Meghir, Narita, and Robin (2015); Charlot, Malherbet, and Ulus (2013); Bosch and Esteban-Pretel (2012); Albrecht, Navarro, and Vroman (2009)]
- ▶ Literature that estimate (or calibrate) a search model and strongly differentiate between the two states.  
[Bobba, Flabbi, and Levy (2022); Bobba, Flabbi, Levy, and Tejada (2021); Narita (2020)]
- ▶ Papers beyond the search literature that take a stand in the debate.  
[Esteban-Pretel and Kitao (2021); Ulyssea (2018); Granda and Hamann (2015); Almeida and Carneiro (2012)]

## Data description

**We use data for Colombia:** A country with high levels of informality and good quality data.

**Gran Encuesta Integrada de Hogares (GEIH):** Monthly household survey that focuses on labor market outcomes.

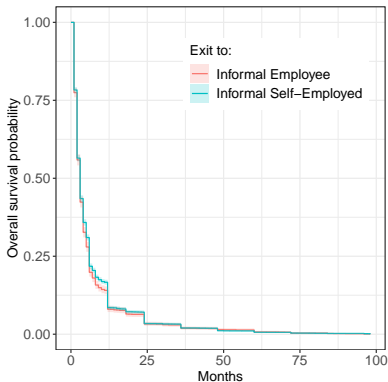
- ▶ Individual characteristics (gender, age, years of schooling)
- ▶ Labor market states:
  - ▶ Unemployed.
  - ▶ Formal employment (full-time employees who contribute to the social security).
  - ▶ Informal employment (full-time informal employees + self-employed working 48+ hours a week (top coded at 100 hours)).
- ▶ Ongoing durations in unemployment and employment states (in months).
- ▶ Labor income and weekly hours worked:
  - ▶ Real monthly wages (in US dollars of December 2016).

## Descriptive statistics on labor market outcomes

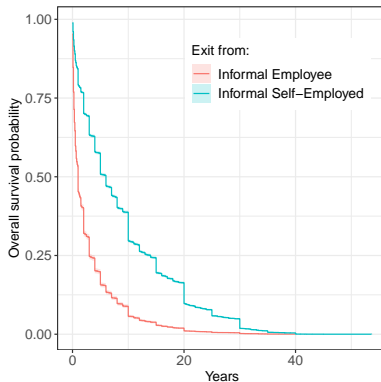
	Unemployed $v = u$	Formal $v = f$	Informal E $v = i$	Informal SE $v = s$
Duration (months):				
$E[t v]$	4.0	64.8	32.8	105.7
$SD[t v]$	6.9	76.6	53.1	103.9
Labor Income (US dollars per hour):				
$E[x v]$	—	1.42	1.06	1.07
$SD[x v]$	—	0.55	0.36	0.56
Labor Market States (Proportion of the Labor Force):				
$v$	0.09	0.40	0.12	0.39

NOTE: The sample consists of 25-55-year-old men living in urban areas, who have completed at most secondary education and work full-time when employed

# Kaplan-Meier Survival Functions

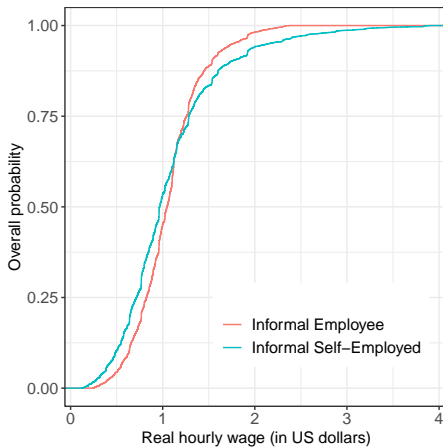


(a) Unemployment Duration conditional on Exit



(b) Employment Duration

## Empirical CDF of accepted hourly wages





# Kolmogorov-Smirnov Test of Equality of Distributions

	Statistic	P-value
$H_0$ : Informal E duration distribution = Informal SE duration distribution:		
Values	0.3869	0.0000
$H_0$ : U duration distribution before Informal E = U duration distribution before Informal SE		
Values	0.0326	0.0177
$H_0$ : Informal E labor income distribution = Informal SE labor income distribution		
Values	0.1106	0.0000

NOTE: E denotes employees, SE self-employed, and U unemployed.

## The model

- We use a four-state partial equilibrium search model: unemployment  $u$ , informal self-employment  $s$ , formal employment  $f$ , and informal employment  $i$ .
- Let  $U$  and  $E(x, v)$  be the steady-state values of unemployment and employment.

$$\rho U = b + \sum_{v=s, f, i} \lambda(v) \left[ \int \max \{E(x, v) - U, 0\} dG(x|v) \right]$$

$$\rho E(x, v) = x [1 - \tau \iota_{v=f}] + \theta \iota_{v=f} + \eta(v) [U - E(x, v)]$$

- The reservation labor income  $x^*(v) = \frac{\rho U - \theta \iota_{v=f}}{1 - \tau \iota_{v=f}}$ , with  $\theta = \tau \rho U$ .
- Steady-state equilibrium:  $u \lambda(v) [1 - G(x^*(v)|v)] = \eta(v)v$ .
- The *null hypothesis* of behaviorally indistinguishable informal states is:

$$\begin{cases} \lambda(s) = \lambda(i) \\ \eta(s) = \eta(i) \\ G(x|s) = G(x|i) \end{cases}$$

## Estimation

- ▶ We estimate the model by maximum likelihood. The null hypothesis is directly tested by using the LR test.
- ▶ The identification strategy follow standard arguments:
  - ▶ Durations provide direct information to identify hazard rates.
  - ▶ Labor incomes identify wage offers distributions under recoverable parametric distribution [Flinn and Heckman (1982)]: we assume that  $G(x|v)$  are log-normal with parameters  $\{\mu(v), \sigma(v)\}$ .
  - ▶ The flow utility of unemployment ( $b$ ) and the discount rate ( $\rho$ ) are jointly identified. We set  $\rho = 0.12$  [Moore, et.al. (2020)] to recover  $b$ .
  - ▶ We set  $\tau = 0.16$  [Fernandez and Villar (2017)] and recover  $\theta$  from condition  $\theta = \tau\rho U$ .
- ▶ To account for the possibility of labor income being measured with error, we incorporate measurement error in the estimation procedure:  $x^o = x \times \epsilon$

## Estimated parameters

Parameters	Unrestricted			Restricted		
	Formal $v = f$	Informal E $v = i$	Informal SE $v = s$	Formal $v = f$	Informal E $v = i$	Informal SE $v = s$
$\lambda(v)$	0.0897 (0.0015)	0.0553 (0.0012)	0.0529 (0.0009)	0.097 (0.0015)	0.0446 (0.0007)	
$\eta(v)$	0.0157 (0.0001)	0.0317 (0.0006)	0.0096 (0.0001)	0.0159 (0.0001)	0.0115 (0.0001)	
$\mu(v)$	0.2956 (0.0033)	0.0071 (0.0055)	-0.0500 (0.0001)	0.2957 (0.0068)	-0.0363 (0.0068)	
$\sigma(v)$	0.3145 (0.0099)	0.3434 (0.0098)	0.5048 (0.0058)	0.3141 (0.0222)	0.4723 (0.0142)	
$b$		-1.7205 (0.0839)			-1.6053 (0.0954)	
$\theta$		0.0106 (0.0048)			0.0160 (0.0060)	
Log-Likelihood		-470885.0			-480498.0	
LR Statistic		—			19226.0	
P-value		—			0.0000	

NOTE: Bootstrapped standard errors in parenthesis. E denotes employees, and SE self-employed.

# Implied labor market outcomes

Values	Unrestricted			Restricted		
	Formal $v = f$	Informal E $v = i$	Informal SE $v = s$	Formal $v = f$	Informal E $v = i$	Informal SE $v = s$
Employment:						
$E[t v]$	63.5	31.5	104.4	63.0	86.8	
$E[x v]$	1.412	1.068	1.080	1.412	1.078	
$SD[x v]$	0.207	0.143	0.339	0.207	0.291	
Unemployment:						
$E[t u]$		5.1			5.4	

## Concluding remarks

- ▶ The paper performs both a parametric and non-parametric analysis to address a fundamental question in the growing literature using search models to study labor market informality
- ▶ Both analyses strongly reject the equality of the two labor market states, cautioning against aggregating them in a common “informality state”.
- ▶ The sources of the difference are the high dispersion of informal self-employment income offers and the low duration of informal employee jobs.