

# Informality and Labor Market Mobility in Turkey

Evidence from Micro Data, 2000-2002<sup>1</sup>

Osman Zeki Gökçe<sup>†</sup>   İnsan Tunalı<sup>‡</sup>

Turkey Labor Market Network Meeting

---

<sup>1</sup>Earlier version of this study was presented at the ERF 20<sup>th</sup> Annual Conference, 23-25 March 2014, Cairo, Egypt. We thank all the participants and the discussant in this event for helpful feedback. We also would like to thank TÜBİTAK, for supporting us financially, and TURKSTAT, for granting us access to the data set.

<sup>†</sup>Ph.D. Student and Research Assistant at Sabancı University, Political Science, e-mail: ogokce@sabanciuniv.edu

<sup>‡</sup>Associate Professor at Koç University, Economics, e-mail: itunali@ku.edu.tr

# Why does informality matter?

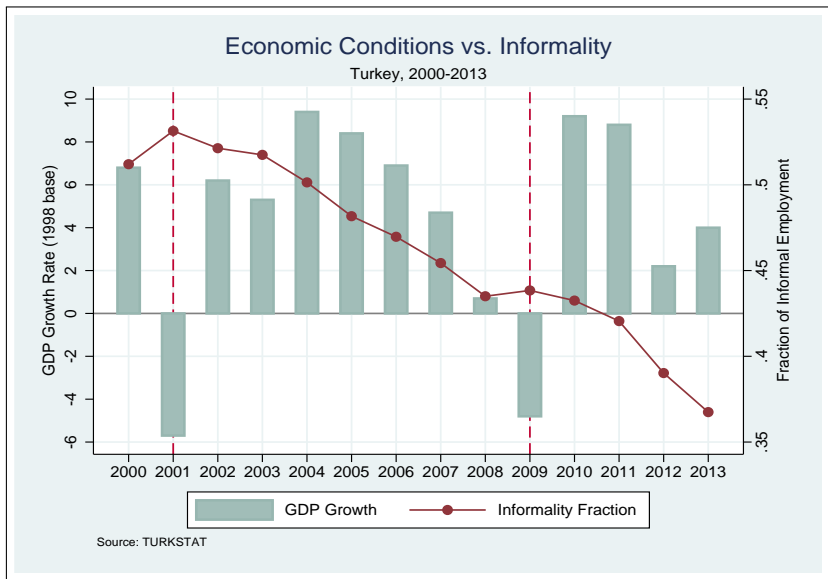
- Informality  $\approx$  low pay, little or no job protection, tax evasion, and sub-standard work conditions (Todaro, 1969; Castells and Portes, 1989)
  - ▶ Unfavorable economic and social consequences for development (Karakurum-Özdemir and Thomas, 2010)
- Involuntary vs. Voluntary
  - ▶ Involuntary (exclusion): Alternative for people who have no other option (Harris and Todaro, 1970)
  - ▶ Voluntary: Based on implicit cost-benefit analyses people may opt for informal employment (Hirschman, 1970)
  - ▶ Perry *et al.* (2007): Both the involuntary and voluntary mechanisms exist

# How to measure informality

- Wide-ranging definitions of informality
  - ▶ Firm level: Firm non-registration or firm size
  - ▶ Individual level:
    - ▶ Occupation type
    - ▶ **Lack of social security coverage**
- We are interested in individuals. Moreover:
  - ▶ Availability of household data makes empirical study reliable (Taymaz, 2009a)
  - ▶ Definitions based on occupation and firm size seem “*arbitrary in practice* even if conceptually well-founded” (Henley *et al.*, 2006)
  - ▶ In Turkey, informality as lack of social security coverage is substantial (Oviedo, 2009)

Therefore, we rely on information on social security coverage.

# Trend of Informality in Turkey



# What does this trend say?

1. Conspicuous decline in fraction of informality: Is sustained economic growth enough to explain this decline?
  - ▶ Government policies against informality
  - ▶ Decline in the share of agriculture in employment (İlkkaracan and Tunalı, 2010)
2. When fortunes reverse, informality share goes up

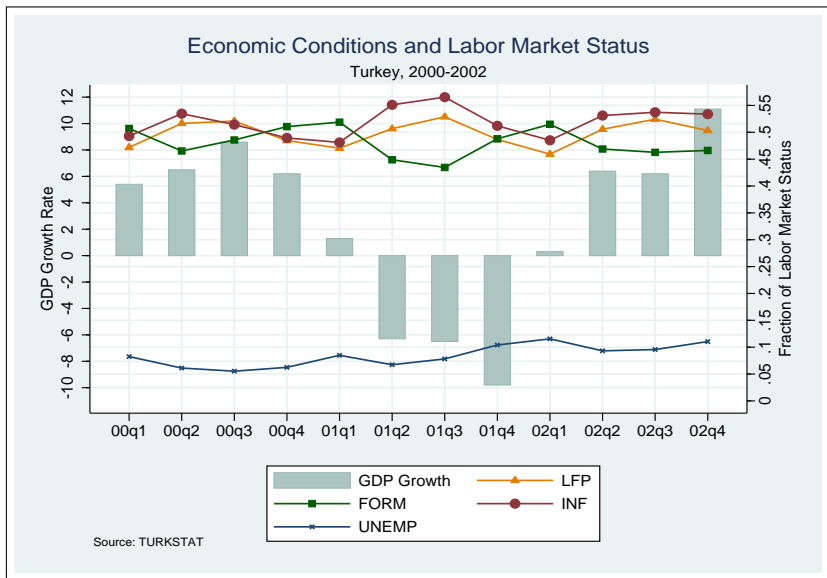
## Drawback of using aggregate data

- However, aggregate data do not inform us effectively about individual labor market experiences over time
  - ▶ We do not know whether individuals cross the boundary and switch employment states
  - ▶ If so, who is more likely to switch states
- We study the banking crisis period of Turkey (2000-2002)<sup>4</sup> and contribute to this debate by obtaining *unbiased* estimates of the magnitudes of flows into/out of, and between Informal and Formal employment

---

<sup>4</sup>2000 is neither a crisis nor a recovery year. 2001 is the year when crisis erupted. 2002 might be seen as a recovery year.

# Turkey, 2000-2002



# How share of informal employment adjusts

- Counter-cyclical:
  - ▶ Application of the Harris and Todaro (1969) model: Due to downward rigidity of wages in the formal sector, the formal sector has difficulties in adjusting to adverse shocks, and the informal sector absorbs workers who would otherwise be unemployed (Maloney, 1999; Loayza and Rigolini, 2006)
- Pro-cyclical:
  - ▶ Finess, Fugazza, and Maloney (2006): A boom in a non-tradable sector such as construction opens new work opportunities for informal contractors and employees, and could cause a pro-cyclical expansion of the sector

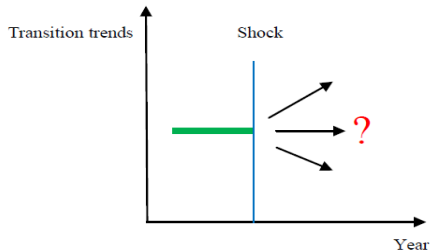


# How mobility between formal and informal employment adjusts

- Bosch, Goni, and Maloney (2006): Mobility falls in downturns, and increases in upturns
  - ▶ Downturns  $\Rightarrow$  consolidation of existing positions
  - ▶ Upturns  $\Rightarrow$  new opportunities  $\Rightarrow$  increase in mobility
- Bosch and Maloney (2006):
  - ▶  $P(\text{INF} \rightarrow \text{FORM}) > P(\text{FORM} \rightarrow \text{INF})$
  - ▶  $P(\text{in/outflows to/from INF})$  is more volatile
  - ▶  $P(\text{INF} \rightarrow \text{UNEMP}) > P(\text{FORM} \rightarrow \text{UNEMP})$  in crisis times

# What will we do basically?

- We estimate:



- ▶ Forward transitions probabilities to see trends in flows over 2000-2002
- ▶ Multinomial Logit Models on micro data to test claims on cyclicity over 2000-2002
  - ▶ Time dummies: The causal effect of shocks + Controls: Sex, location, age, education..
  - ▶ Question: Holding sex, location, age, education.. constant, does transition behavior change when a shock hits the economy?

# Data:

- The Turkish Household Labor Force Survey (HLFS), 2000-2002
- In every period, HLFS includes about 70,000 individuals from 18,000 - 20,000 households
- Information on demography and labor market statuses of individuals
- Short panel component  $\Rightarrow$  dynamic analyses of transition rates
  - ▶ 11 quarterly panels (i.e. 2000: Q1-Q2, Q2-Q3,..., 2001: Q1-Q2, Q2-Q3,.. and 2002: Q1-Q2, Q2-Q3,...)
  - ▶ 8 annual panels (Q1: 2000-2001 and 2001-2002, Q2: 2000-2001 and 2001-2002,...)

## Biases on the data:

- Tunalı, Ekinci, and Yavuzoğlu (2012): Due to address-based rotating sampling frame, a short panel component suffers from:
  - ▶ Attrition: Initial response followed by non-response:
    - ▶ Tunalı (2009) establishes that attrition observed in HLFS during the 2000-2002 period is severe
  - ▶ Substitution: Non-response followed by response
- We will use the RAN model to correct for attrition and substitution suggested by Tunalı, Ekinci, and Yavuzoğlu (2012)

- Step 1: RAN Model to correct attrition and substitution
- Step 2: MNL Estimation to test determinants of forward transitions using corrected data

Consider a two-round panel and let;

$y_{ij}$  = the labor market status of individual  $i$  in period  $j$ ,

$x_i$  = the observed characteristics (education level, age, sex, etc.) of individual  $i$  in the first period,

$D_i = 1$  if individual  $i$  responds in both periods, 0 else.

Object of interest:

$f(y_1, y_2|x)$ : the joint distribution of labor market states, conditional on  $x$ .

Key question: Whether  $P(D = 1|y_1, y_2, x) = P(D = 1|x)$ , or not?

- If yes, selection into  $D = 1$  is 'ignorable'
- However, Tunalı (2009): attrition process is influenced by the labor market states.

$\Rightarrow f(y_1, y_2|x, D = 1) \neq f(y_1, y_2|x)$

RAN key equation:

$$f(y_1, y_2|x) = \frac{f(y_1, y_2|x, D = 1) P(D = 1|x)}{P(D = 1|y_1, y_2, x)} \quad (1)$$

The resulting RAN model, which rescales the retention probability, can be expressed as:

$$f(y_1, y_2|x) = f(y_1, y_2|x, D = 1) w(y_1, y_2|x) \quad (2)$$

## Strategy: Choosing functional form of $w(\cdot)$

4x4 RAN Model: NP(=0), INF(=1), FORM(=2), UNEMP(=3). We treat non-participation in both periods as the reference category, and introduce 6 indicators (3 per period):

$$\begin{aligned}
 z_{1j} &= \begin{cases} 1, & \text{employed in the informal sector}(y_j = 1) \\ 0, & \text{otherwise} \end{cases} \\
 z_{2j} &= \begin{cases} 1, & \text{employed in the formal sector}(y_j = 2) \\ 0, & \text{otherwise} \end{cases} \\
 z_{3j} &= \begin{cases} 1, & \text{unemployed}(y_j = 3) \\ 0, & \text{otherwise} \end{cases}
 \end{aligned} \tag{3}$$

We rely on linear refraction function:

$$w(\underline{z}_1, \underline{z}_2) = \theta_0 + \theta_1 z_{11} + \theta_2 z_{21} + \theta_3 z_{31} + \theta_4 z_{12} + \theta_5 z_{22} + \theta_6 z_{32}. \tag{4}$$

We use MATLAB to estimate  $\theta$ .

Refraction function captures the propensity to be absent in the balanced panel as a function of labor market states occupied in periods 1 and 2.

Identifying information comes from marginals published by TURKSTAT:

$$\sum_{y_2} f(y_1, y_2 | x, D = 1) w(\underline{\theta}' \underline{z} | x) = f_1(y_1 | x), \quad (5)$$

$$\sum_{y_1} f(y_1, y_2 | x, D = 1) w(\underline{\theta}' \underline{z} | x) = f_2(y_2 | x). \quad (6)$$

In the RAN Model dependence of refutation factors on;

- $(y_1, y_2)$  is parametric,
- $x$  is non-parametric

Choice of  $x$ :

All (age 15+)

By sex: Male, female

By sex\*location: Urban males, rural males, urban females, rural females



## Tabular representation of the 4x4 problem:

DATA:

$P_{y_1, y_2} = f(y_1, y_2 | D=1)$ , fractions in the balanced panel.

$f_1(y_1)$  and  $f_2(y_2)$ , “unbiased” marginals (published by TURKSAT).

	$y_2=0$	$y_2=1$	$y_2=2$	$y_2=3$	
$y_1=0$	$\theta_0 P_{00}$	$(\theta_0 + \theta_2) P_{01}$	$(\theta_0 + \theta_4) P_{02}$	$(\theta_0 + \theta_6) P_{03}$	$f_1(0)$
$y_1=1$	$(\theta_0 + \theta_1) P_{10}$	$(\theta_0 + \theta_1 + \theta_2) P_{11}$	$(\theta_0 + \theta_1 + \theta_4) P_{12}$	$(\theta_0 + \theta_1 + \theta_6) P_{13}$	$f_1(1)$
$y_1=2$	$(\theta_0 + \theta_3) P_{20}$	$(\theta_0 + \theta_3 + \theta_2) P_{21}$	$(\theta_0 + \theta_3 + \theta_4) P_{22}$	$(\theta_0 + \theta_3 + \theta_6) P_{23}$	$f_1(2)$
$y_1=3$	$(\theta_0 + \theta_5) P_{30}$	$(\theta_0 + \theta_5 + \theta_2) P_{31}$	$(\theta_0 + \theta_5 + \theta_4) P_{32}$	$(\theta_0 + \theta_5 + \theta_6) P_{33}$	$f_1(3)$
	$f_2(0)$	$f_2(1)$	$f_2(2)$	$f_2(3)$	

Objective: Choose  $\Theta = \{\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$  so that row & column restrictions are met.

We have 7 linearly independent equations in 7 unknowns.

We express the “reflation factors”  $w(y_1, y_2|x)$  as a function of  $y_1$  and  $y_2$ .

$$w(y_1, y_2|x) = 1 : \text{no bias,} \quad (7)$$

$$w(y_1, y_2|x) > 1 : \text{“downward bias” or “under-represented” in BP,} \quad (8)$$

$$w(y_1, y_2|x) < 1 : \text{“upward bias” or “over-represented” in BP} \quad (9)$$

## Annual, 2001-Q1 to 2002-Q1, 15+ all

Bold items; obtained directly from available data.

Joint probabilities; come from the balanced panel component of the individual level data.

Marginals; come from TURKSTAT web site.

The reflation factors (non-bold items); estimated by using MATLAB.

	$y_2 = 0$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$	Margins	TURKSTAT
$y_1 = 0$	0.8923 x <b>0.5052</b>	1.1407 x <b>0.0388</b>	0.9723 x <b>0.0177</b>	1.0714 x <b>0.0159</b>	<b>0.5778</b>	$f_1(0)$ 0.5295
$y_1 = 1$	1.1272 x <b>0.0471</b>	1.3756 x <b>0.0849</b>	1.2072 x <b>0.0197</b>	1.3063 x <b>0.0101</b>	<b>0.1620</b>	$f_1(1)$ 0.2071
$y_1 = 2$	0.8862 x <b>0.0268</b>	1.1346 x <b>0.0235</b>	0.9663 x <b>0.1669</b>	1.0653 x <b>0.0108</b>	<b>0.2282</b>	$f_1(2)$ 0.2233
$y_1 = 3$	1.1355 x <b>0.0112</b>	1.3839 x <b>0.0066</b>	1.2155 x <b>0.0055</b>	1.3146 x <b>0.0085</b>	<b>0.032</b>	$f_1(3)$ 0.0399
Margins	<b>0.5905</b>	<b>0.1540</b>	<b>0.21</b>	<b>0.0454</b>	<b>1</b>	
TURKSTAT	$f_2(0)$ 0.5406	$f_2(1)$ 0.1971	$f_2(2)$ 0.2092	$f_2(3)$ 0.0530		1

## Statistical question:

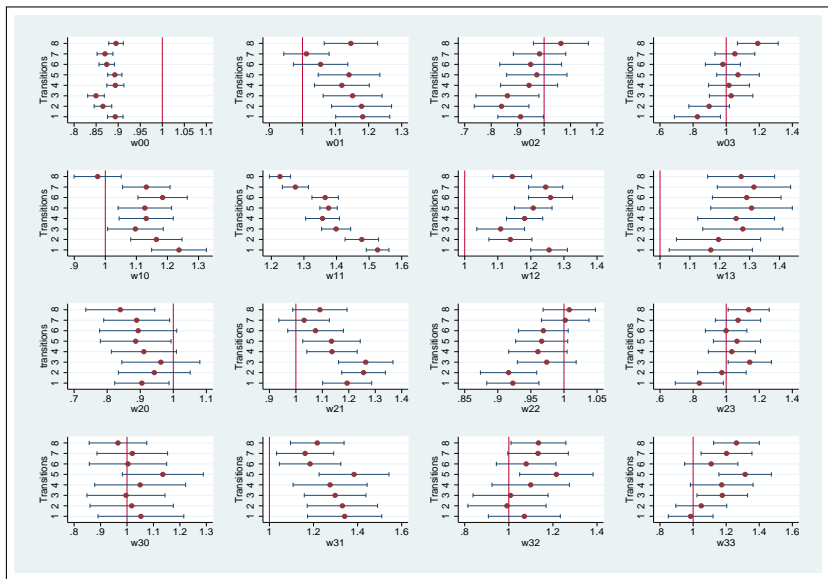
### Is attrition ignorable?

We should test  $H_0 : w(y_1, y_2|x) = 1$  vs.  $H_1 : w(y_1, y_2|x) \neq 1$ .

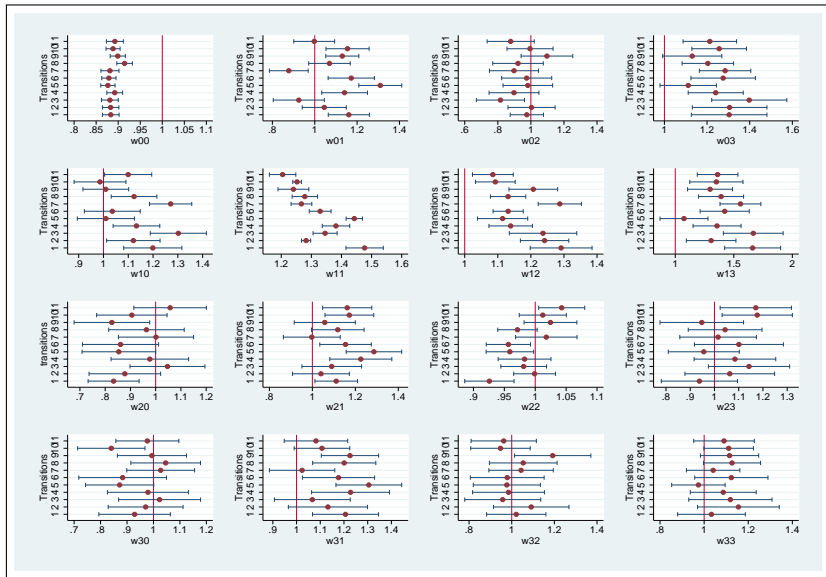
Attrition has economic implications as well. Recall Bosch and Maloney (2006): P(in/outflows to/from INF) is more volatile. Thus:

**Hypothesis 1:** *Compared to others, individuals who are in the informal state are more likely to be attritors/substitutes (downward bias or under-representation in the BP).*

# Weights for annual transitions:



# Weights for quarterly transitions:



# Attrition is not ignorable!

$$H_0 : w(y_1, y_2|x) = 1 \text{ vs. } H_1 : w(y_1, y_2|x) \neq 1.$$

Clear rejection of the null: Need the weights

**Hypothesis 1:** *Compared to others, individuals who are in the informal state are more likely to be attritors/substitutes (downward bias or under-representation in the BP).*

Hypothesis 1: Corroborated by the data

The probability of transition from state  $m$  to state  $n$  (where  $m, n = 0, 1, 2, 3$ ) is given by:

$$Prob(y_{i1} = n | y_{i2} = m) = \frac{\exp(\underline{x}'_i \underline{\beta}^n)}{1 + \sum_{m=1}^3 \exp(\underline{x}'_i \underline{\beta}^m)}, \quad n = 1, 2, 3; \quad (10)$$

$$Prob(y_{i1} = m | y_{i2} = m) = \frac{1}{1 + \sum_{m=1}^3 \exp(\underline{x}'_i \underline{\beta}^m)} \quad (11)$$

- 4 states: NP, INF, FORM, and UNEMP.
- Reference: Individuals retaining their positions across both periods.
- Using the MNL model we can determine whether a particular group of individuals identified by their observed characteristics is relatively more likely to end up in a particular destination state after one-year (or one-quarter).



Our explanatory variables are:

- Variables to see the effects of economic shock:
  - ▶ year dummies (Year 2000, Year 2001, Year 2002, Year 2001/2002)
  - ▶ unemployment rates based on location (rural or urban) (Unemployment rate)
- Control variables:
  - ▶ seasonality (Quarter 1, Quarter 2, Quarter 3, Quarter 4)
  - ▶ gender\*location (Urban-Male, Urban-Female, Rural-Male, Rural-Female)
  - ▶ age (15-24, 25-34, 35-44, 45-54, 55-64 and 65+)
  - ▶ education levels (Illiterate, Literate but no diploma, Primary 5 years, Primary 8 years, Middle School, General High School, Vocational High School, Associate (less than 4 years) of higher education, University (4 years or more))
  - ▶ marital status (Single, Divorced, Married, Widow)
  - ▶ in-family status of individual (Household Head, Non-Household Head)

Reference category: urban-male, ages between 35 and 44, 5 years of primary education, married, non-household head, into quarter 1 of year 2000.

Both weighted (by  $w(y_1, y_2|x)$ ) and unweighted MNL Estimates are obtained.

# Hypotheses:

**Hypothesis 2:** *(Cyclicalitity of flows) As economy turn upwards, while  $INF \rightarrow FORM$  becomes more likely,  $FORM \rightarrow INF$  becomes less likely.*

**Hypothesis 3:** *(Consolidation of existing position) The higher the unemployment rate, the lower the flows out of  $INF$  and  $FORM$ , in other words out of being employed.*

# MNL estimates: Annual and quarterly

	Transitions from INF(t-1) Base category: INF(t)						Transitions from FORM(t-1) Base category: FORM(t)					
	ANNUAL FORWARD TRANSITION ESTIMATES											
	NP(t)		FORM(t)		UNEMP(t)		NP(t)		INF(t)		UNEMP(t)	
	BP	WP	BP	WP	BP	WP	BP	WP	BP	WP	BP	WP
<b>1<sup>st</sup> Visit: 2001</b> (Base: "1st Visit: 2000")	-0.0011 (0.057)	0.11 (0.058)	0.32*** (0.089)	0.55*** (0.090)	-0.12 (0.15)	-0.040 (0.15)	-0.065 (0.14)	-0.26 (0.14)	-0.23* (0.095)	-0.47*** (0.094)	-0.055 (0.22)	-0.20 (0.22)
<b>Unemployment Rate</b>	-0.025 (0.030)	-0.081** (0.030)	-0.18*** (0.039)	-0.26*** (0.039)	0.12* (0.060)	0.12 (0.060)	-0.0092 (0.052)	0.029 (0.053)	0.088* (0.039)	0.16*** (0.039)	0.085 (0.084)	0.14 (0.085)
	QUARTERLY FORWARD TRANSITION ESTIMATES											
	NP(t)		FORM(t)		UNEMP(t)		NP(t)		INF(t)		UNEMP(t)	
	BP	WP	BP	WP	BP	WP	BP	WP	BP	WP	BP	WP
<b>Quarterly Transitions in 2001</b>	0.027 (0.038)	0.016 (0.038)	-0.0089 (0.048)	-0.069 (0.048)	0.62*** (0.082)	0.50*** (0.086)	-0.16* (0.068)	-0.094 (0.069)	0.10* (0.047)	0.14** (0.048)	0.60*** (0.10)	0.53*** (0.10)
<b>Quarterly Transitions in 2002</b> (Base: "Quarterly Transitions in 2000")	-0.24*** (0.056)	-0.25*** (0.057)	0.069 (0.076)	0.0020 (0.078)	0.68*** (0.13)	0.64*** (0.14)	-0.36** (0.11)	-0.33** (0.11)	0.17* (0.076)	0.21** (0.077)	0.53** (0.18)	0.53** (0.18)
<b>Unemployment Rate</b>	0.053*** (0.013)	0.053*** (0.013)	-0.018 (0.015)	-0.012 (0.015)	-0.028 (0.024)	-0.028 (0.025)	0.033 (0.021)	0.0090 (0.021)	-0.016 (0.015)	-0.017 (0.015)	-0.054 (0.031)	-0.055 (0.031)

## Test results:

- Evidently weighted vs. balance differences are substantial in the case of year dummies.

**Hypothesis 2:** (*Cyclical flow of flows*) As economy turn upwards, while  $INF \rightarrow FORM$  becomes more likely,  $FORM \rightarrow INF$  becomes less likely.

- Hypothesis 2: Verified for yearly transitions

**Hypothesis 3:** (*Consolidation of existing position*) The higher the unemployment rate, the lower the flows out of  $INF$  and  $FORM$ , in other words out of being employed.

- Hypothesis 3: Not as claimed, but we could not falsify it either (again, in yearly flows)

## To wrap up:

- Selection in HLFS is **not ignorable and has economic implications**.
  - ▶ Compared to others, individuals who are in the informal state are more likely to be attritors/substitutes
- There is **cyclical flow** (in yearly transitions)
- We observe somewhat **consolidation of existing position** during unfavorable conditions (in yearly transitions)

THANK YOU!