The Chicken or the Egg?
Informality, Productivity and Economic Growth in Mexico

Columbia University
April, 2015.

Matias Busso and Santiago Levy,
Inter-American Development Bank.
Contents

1. Motivation
2. Simple framework
3. Data and numerical results
4. Disentangling "distortions" in $\mathcal{E}(.)$
5. Final observations
1. Motivation
In the 1990-2013 period:

• Average annual growth rate of per capita GDP of only 1.3% despite:

  ✓ successful macro stabilization after the 1994-95 crisis
  ✓ large trade reforms (Nafta and other FTAs):
    ◦ increase in exports from 5 to 25% of GDP, and
    ◦ share of manufactures in merchandise exports raised from 40 to 80%
  ✓ increase of 42% in years of schooling of labor force (2.9 years)

• Stagnant TFP growth:

  ✓ -0.53% annual average rate in \( Y = A.F(K,hL) \) with \( h > 1 \) capturing human capital, or
  ✓ +0.14% annual average rate in \( Y = A*.F(K,L) \) including human capital in \( A* \)

• Persistent informality:

  ✓ share of labor force informally employed constant at 60%, and
  ✓ share of informal firms at 75% also constant (data only from 1998 onwards)
We study the relation between low productivity and informality:

- Is low TFP growth the result of a large informal sector? If so, what causes so much informality?
- Is large informality the result of low TFP growth? If so, what lies behind low TFP growth?
- Or are low TFP growth and large informality the result of some other underlying characteristic of the economy? If so, is that because of “distortions” or “low human capital”?

Issue relevant for other countries in LA also characterized by high informality and low productivity growth.
2. Simple framework
“Distortions” (broadly) and human capital

• The “Environment”, \( \mathcal{E} = \{L, C, T, Z\} \)

  \( L = \) {regulations on social insurance for salaried and non-salaried labor (coverage, bundling and financing); regulations on salaried labor regarding length of contracts, minimum wages, hiring, firing, promotion, unionization and the like; workers’ valuations of social insurance benefits; enforcement of regulations}

  \( C = \) {costs and conditions of access to credit from private banks, venture capital or other sources (rates, terms, collateral); access to public development banks; micro-credit programs; regulations on bankruptcy}

  \( T = \) {income taxes on firms; taxes on salaried and non-salaried earnings; value added taxes; special tax regimes (by firm size or sector); costs of compliance; enforcement of tax regulations}

  \( Z = \) {trust in labor, tax and civil courts; institutions for contract enforcement and for protection of property and intellectual rights; other public goods; registration costs; trade regime; product market competition; subsidies to activities like R&D; infrastructure, ......}

• \( \mathcal{E}(.) \) reflects elements of public policy (including the functioning of institutions), any market failures, and some aspects of individual preferences and beliefs (attitudes towards risks, valuation of social insurance benefits, trust). It is the context in which individuals make decisions.
Individuals in the economy can be workers (salaried or non-salaried), self-employed, or entrepreneurs (with workers in a firm). Individuals have assets which may be used for production (land, street corner, truck, machine), and in some cases also pledged as collateral.

Each individual has effective labor l(s,a) where s captures schooling (years/quality) and a abilities (innate or acquired on the job).

Individuals may have access to credit for production or investment given regulations of C in E(.,C,.,.), and assets that they can pledge as collateral.

Factors of production set is F[l(s,a), k] where k is physical capital. (But here we focus on human capital.)

At any point in time F[l(s,a), k] is given, although human capital can be produced by investments in early child development and schooling; or acquired by workers through training or on-the-job learning. (Physical capital can be produced by investments.)

“Technology set” T contains best practices and, importantly, whether for any activity salaried vs. non-salaried contractual arrangements are preferred (monitoring and coordination of tasks, vs. risk or profit sharing and effort eliciting).
Realizations (or equilibrium outcomes)

• Everybody maximizes given the regulatory environment:

\{ E(L,C,T,Z), T, F[l(s,a), k] \} \rightarrow R, \text{ where } R \text{ provides a complete picture of the economy}

✓ who is an entrepreneur (and implicitly how many firms),

✓ who is self-employed,

✓ who is a worker (salaried or non-salaried)

✓ size and contractual structure of each firm (number of workers and salaried/non-salaried mix)

✓ degree of compliance with various regulations of L, C, T and Z in E (or margins of illegality in dimensions like paying taxes, observing minimum wage laws, and so on)

✓ returns to physical capital and human capital (education, skills)

✓ .........
Misallocation and observed and potential productivity

- Optimal environment $\mathbf{E}^*$ occurs when regulation is perfect: $L = L^*$, $C = C^*$, $T = T^*$ and $Z = Z^*$, and the economy produces with maximum efficiency.

$$R \{\mathbf{E}^*(L^*,C^*,T^*,Z^*), T, F[l(s,a),k]\} \rightarrow \mathbf{P}^*,$$

where $\mathbf{P}$ is an index of aggregate productivity, and an * denotes optimal values.

- When $\mathbf{E} = \mathbf{E}^*$, $MRPL_{ij}$ same for all sectors $i$ and firms $j$; similarly for $MRPK_{ij}$.

- $\mathbf{E}$ differs from $\mathbf{E}^*$ as a result of “distortions” (interpreted very broadly, as policy or market failures that impede channeling resources to their most productive use).

- “Distortions” generate misallocation, so observed $\mathbf{P} \{\mathbf{E}, T, F\} < \mathbf{P}^* \{\mathbf{E}^*, T, F\}$

Note that:

- two types of misallocation may occur: (i) along extensive margin, between individuals as to who is an entrepreneur, who self-employed and who a worker, and (ii) along intensive margin, as existing firms (= entrepreneurs) use the wrong amounts of $k$ and $l$.

- “low productivity” is equivalent to misallocation,

- “low productivity” is measured relative to the potential of the economy given by $T$ and $F[l(s,a),k]$, not with respect to another economy.
Informality

• Following Kanbur (2010), informality is defined with respect to a specific regulation. Within E(.) we can chose regulations from L, T and Z (less common to pick from C); precision needed to avoid confusion.

• We choose the following specific L-regulation: “Salaried workers enrolled in contributory social insurance (CSI) by the firms who hire them are formal, with contributions proportional to wages”.

• All other workers are informal: self-employed, non-salaried, or salaried in illegal firms.

• Five firm types:
  ✓ 1. formal and legal, only salaried workers all enrolled in CSI
  ✓ 2. informal and legal, only non-salaried workers
  ✓ 3. informal and illegal, only salaried workers but none enrolled in CSI
  ✓ 4. semi-formal and legal, mixes salaried and non-salaried workers, but all salaried in CSI
  ✓ 5. semi-formal and semi-legal, mixes salaried and non-salaried, but not all salaried in CSI

• Note that:
  ✓ informality and illegality not equivalent
  ✓ “informal sector”, I, not defined with precision; I = all informal workers + type 2 and 3 firms + ......; there is a fuzzy middle given type 4 and type 5 firms.

• Highlight: the distinction between salaried and non-salaried workers central to analysis.
Informality and low productivity

• Informality only observed once realizations $R$ occur, as we see the distribution of individuals between occupations and the type distribution of firms. Similarly, “low productivity” (i.e., $P < P^*$), only observed once realizations $R$ occur and intensive and extensive margins fixed. We have:

$$R \{E, T, F\} \rightarrow P \leftarrow I$$

• Key point: misallocation (= low productivity) and informality jointly determined.

• Given $T$ and $F[l(s,a),k]$, at any point in time the distribution of individuals among occupations, the size and type distribution of firms and of workers across labor contracts, are endogenous to $E$.

“Low productivity” and informality are outcomes of this process.

• However, even if $E = E^*$ and thus, $P = P^*$, need not imply that $I^* = 0$.

• “Some” informality (under our definition) is consistent with $P^*$ (non-salaried contracts necessary in contexts where effort is not observed or risk needs to be shared; some self-employment optimal).

• When $E$ differs from $E^*$, then $P < P^*$, but not clear whether $I < I^*$; this depends on the nature of “distortions”. (If non-salaried contracts are taxed, there may be “too little” informality.)
Four empirical questions

• Q1: Does $E$ differ from $E^*$, and therefore $P < P^*$? Since $E$ and $E^*$ are not observed, we test this indirectly by measuring differences in the MRP of K and L of firms in same sector.

• Q2: Are “distortions” causing too much or too little informality? We test this by comparing the MRP of K and L across firm types.

• Q3: Given $E$, will an increase in $F[l(s,a)]$ raise $P$? We test this by comparing the MRP of K & L of firms of the same type in two economies with the same $E$ but different $F$.

• Q4: Given $E$, will an increase in $F[l(s,a)]$ decrease informality? We test this by comparing the type distribution of firms in two economies with the same $E$ but different $F$.

• We use these questions to try to disentangle the effects of “distortions” and human capital on productivity, on one hand; and informality, on the other.

• We separately focus on the relation between firm size, firm type and productivity.

• We use Hsieh and Klenow’s (2009) framework to measure the physical (TFPQ) and revenue (TFPR) productivities of firms.

- H-K give more structure to the mapping \( \mathbb{R} \{ E(L,C,T,Z), T, F[l(s,a), k] \} \rightarrow P \)

- **T**: Cobb-Douglas with CRS and capital coefficient \( \alpha_s \) equal for all firms \( i \) in sector \( s \)

- **F**: \( l(s,a) \) wage bill; \( k \) book value of capital stock

- **E**: “distortions” captured through idiosyncratic wedges in output \( \tau_{Q_{is}} \) and factor prices \( \tau_{L_{ls}} \) such that firm’s maximization problem is:

\[
\pi_{is} = (1 - \tau_{Q_{is}})P_{is}Q_{is} - (1 - \tau_{L_{is}})wL_{is} - rK_{is} \quad \text{where} \quad Q_{is} = A_{is}K_{is}^{\alpha_s}L_{is}^{1-\alpha_s}
\]

- Each firm \( i \) in sector \( s \) faces a negatively sloped demand curve as they produce an imperfect substitute to goods produced by other firms in the same sector \( s \)

- Key distinction between the **physical productivity** of each firm, \( \text{TFPQ}_{is} = A_{is} \), and the **revenue productivity**, \( \text{TFPR}_{is} \), which is equal to: \( \text{PisA}_{is} \)
3. Data and numerical results
Data

• Economic Census for 1998, 2003 and 2008, which contain detailed firm (= establishment) level data on value added, capital, labor, social insurance contributions, and so on.

• Census captures firms of all sizes in fixed establishments in urban areas; rural activity and urban activity in the streets excluded. Data disaggregated at state, municipal and AGEB level.

• Unfortunately, no data on education and human capital. However, Population Census with same level of geographical disaggregation does have data on years of schooling.

• For 2008, we have data for 3.6 million firms (indexed by i), which are classified into 707 6-digit sectors (indexed by s).

• We classify firms by type and size, and compute each firm’s $\text{TFP}_{is}$ and $\text{TFPR}_{is}$.

• Firms in Census account for about 50% of total employment. Remaining employment occurs in rural areas, the public sector and urban activity not captured in the Census (in streets).

• We estimate that urban workers not captured in the Census are employed in an additional 1.2 million firms, for which we have no data on capital or social insurance contributions. In all likelihood these are all informal firms. Same holds for urban self-employed workers and firms and workers in rural areas. Thus, while our data is exceptionally rich, we fail to fully capture informal activity.
## Summary of descriptive statistics

<table>
<thead>
<tr>
<th>Establishments</th>
<th>Legal &amp; formal</th>
<th>Legal &amp; Informal</th>
<th>Legal &amp; Semi-formal</th>
<th>Semi-legal &amp; Semi-formal</th>
<th>Illegal and informal</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0-5]</td>
<td>1.21</td>
<td>65.60</td>
<td>1.18</td>
<td>2.49</td>
<td>19.25</td>
</tr>
<tr>
<td>[6-10]</td>
<td>0.69</td>
<td>1.11</td>
<td>0.33</td>
<td>1.23</td>
<td>2.46</td>
</tr>
<tr>
<td>[11-50]</td>
<td>0.87</td>
<td>0.56</td>
<td>0.15</td>
<td>1.10</td>
<td>0.90</td>
</tr>
<tr>
<td>[+50]</td>
<td>0.29</td>
<td>0.19</td>
<td>0.03</td>
<td>0.30</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Workers</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[0-5]</td>
<td>0.77</td>
<td>23.60</td>
<td>0.79</td>
<td>1.69</td>
<td>10.94</td>
</tr>
<tr>
<td>[6-10]</td>
<td>1.09</td>
<td>1.63</td>
<td>0.50</td>
<td>1.92</td>
<td>3.66</td>
</tr>
<tr>
<td>[11-50]</td>
<td>3.84</td>
<td>2.39</td>
<td>0.57</td>
<td>4.72</td>
<td>3.36</td>
</tr>
<tr>
<td>[+50]</td>
<td>14.67</td>
<td>7.74</td>
<td>1.19</td>
<td>13.75</td>
<td>1.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Capital</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[0-5]</td>
<td>0.73</td>
<td>7.17</td>
<td>0.64</td>
<td>1.33</td>
<td>3.35</td>
</tr>
<tr>
<td>[6-10]</td>
<td>0.85</td>
<td>0.78</td>
<td>0.29</td>
<td>1.23</td>
<td>1.30</td>
</tr>
<tr>
<td>[11-50]</td>
<td>3.18</td>
<td>2.15</td>
<td>0.35</td>
<td>3.29</td>
<td>1.28</td>
</tr>
<tr>
<td>[+50]</td>
<td>24.31</td>
<td>21.75</td>
<td>1.10</td>
<td>24.59</td>
<td>0.34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value Added</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[0-5]</td>
<td>1.24</td>
<td>4.71</td>
<td>0.26</td>
<td>1.07</td>
<td>3.03</td>
</tr>
<tr>
<td>[6-10]</td>
<td>0.95</td>
<td>0.78</td>
<td>0.20</td>
<td>1.37</td>
<td>1.35</td>
</tr>
<tr>
<td>[11-50]</td>
<td>3.69</td>
<td>2.17</td>
<td>0.33</td>
<td>4.67</td>
<td>1.66</td>
</tr>
<tr>
<td>[+50]</td>
<td>24.85</td>
<td>12.98</td>
<td>1.15</td>
<td>32.83</td>
<td>0.73</td>
</tr>
</tbody>
</table>

The distinction between firm type and size is of the essence:

- more small formal firms than large formal firms
- more legal informal firms than illegal informal firms
Q1: does $\mathcal{E}$ differ from $\mathcal{E}^*$, and therefore $\mathcal{P} < \mathcal{P}^*$?

- Hsieh and Klenow (2009) show that for firm $i$ in sector $s$ revenue productivity is

$$\text{TFPR}_{si} \propto (\text{MRPK}_{si})^{\alpha_s} (\text{MRPL}_{si})^{1-\alpha_s} \propto (1 + \tau_{L_{si}})^{1-\alpha_s} (1 - \tau_{Q_{si}})$$

where $\tau^{'s}$ measure wedges between MRP of K & L resulting from “distortions”.

- Without “distortions” all $\tau^{'s}$ and TFPR is the same for all firms in a sector.

### Dispersion of TFPR, 2008
(6-digit disaggregation)

<table>
<thead>
<tr>
<th></th>
<th>Std. Dev.</th>
<th>p75-p25</th>
<th>p90-p10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate economy</td>
<td>1.18</td>
<td>1.51</td>
<td>2.98</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>1.11</td>
<td>1.43</td>
<td>2.78</td>
</tr>
<tr>
<td>Commerce</td>
<td>1.33</td>
<td>1.77</td>
<td>3.43</td>
</tr>
<tr>
<td>Services</td>
<td>1.18</td>
<td>1.46</td>
<td>3.00</td>
</tr>
</tbody>
</table>

- Note that since we can only use data of existing firms, we only measure misallocation along the intensive margin; also, we only look at misallocation within sectors. Misallocation probably larger.
Gap between $E$ and $E^*$ larger in Mexico

- Using all manufacturing firms larger than 10 employees, Hsieh and Klenow (2009) find that the TFP gains from removing all distortions in the US are 43%.

- Busso, Madrigal and Pages (2010) perform the same exercise imposing the same sample restrictions, and compute TFP gains for: \textbf{Mexico 95\%}, Venezuela 65\%, Bolivia 61\%, El Salvador 61\%, Uruguay 60\%, Argentina 60\%, Ecuador 58\%, Chile 54\%, and Colombia 51\%.

If “distortions” in Mexico’s manufacturing sector equaled those of the US, TFPQ would be 36% higher.

Source: IDB (2010).
Manufacturing sector only, comparison at 4 digit level.
Q2: are distortions causing too much informality?

Log TFPR\(_{si}\)/log average TFPR in sector s
(excluded category: formal and legal)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal semi-formal</td>
<td>-0.312</td>
<td>-0.499</td>
<td>-0.511</td>
<td>-0.220</td>
<td>-0.203</td>
</tr>
<tr>
<td>s.e.</td>
<td>[0.0044]</td>
<td>[0.0068]</td>
<td>[0.0142]</td>
<td>[0.0187]</td>
<td>[0.0361]</td>
</tr>
<tr>
<td>Legal-informal</td>
<td>-0.385</td>
<td>-0.626</td>
<td>-0.420</td>
<td>-0.348</td>
<td>-0.126</td>
</tr>
<tr>
<td>s.e.</td>
<td>[0.0027]</td>
<td>[0.0045]</td>
<td>[0.0127]</td>
<td>[0.0124]</td>
<td>[0.0195]</td>
</tr>
<tr>
<td>Semilegal-semiformal</td>
<td>-0.023</td>
<td>-0.138</td>
<td>-0.074</td>
<td>-0.059</td>
<td>0.018</td>
</tr>
<tr>
<td>s.e.</td>
<td>[0.0025]</td>
<td>[0.0053]</td>
<td>[0.0097]</td>
<td>[0.0097]</td>
<td>[0.0163]</td>
</tr>
<tr>
<td>Illegal informal</td>
<td>-0.153</td>
<td>-0.306</td>
<td>-0.319</td>
<td>-0.241</td>
<td>0.004</td>
</tr>
<tr>
<td>s.e.</td>
<td>[0.0028]</td>
<td>[0.0046]</td>
<td>[0.0094]</td>
<td>[0.0117]</td>
<td>[0.0485]</td>
</tr>
<tr>
<td>Controls for: age state</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>state</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2,244,403</td>
<td>1,975,211</td>
<td>157,088</td>
<td>93,619</td>
<td>18,485</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.048</td>
<td>0.034</td>
<td>0.033</td>
<td>0.025</td>
<td>0.060</td>
</tr>
</tbody>
</table>

- Resources in legal and formal firms between 15 and 38% more valuable than in any other type of firm, particularly informal firms; thus, informal sector too large.

- Resources in legal informal firms always less valuable than in illegal informal firms; composition of informality matters for productivity.

- “Distortions” in \( E(.) \) systematically induce excess resources into firms with non-salaried contracts or with illegal salaried contracts; or create difficulties for resources to go to firms with legal salaried contracts.
Two economies with same $E$, but different $F[l(s,a)]$

<table>
<thead>
<tr>
<th></th>
<th>Country</th>
<th>Mex City</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of schooling</td>
<td>8.05</td>
<td>10.70</td>
</tr>
<tr>
<td>&gt;15 incomplete primary*</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td>&gt;15 complete primary*</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>&gt;18 post-secondary*</td>
<td>0.35</td>
<td>0.56</td>
</tr>
</tbody>
</table>

* shares

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td># of firms</td>
<td>3,643,982</td>
<td>452,102</td>
</tr>
<tr>
<td>Share of K</td>
<td>100</td>
<td>18</td>
</tr>
<tr>
<td>Share of L</td>
<td>100</td>
<td>19</td>
</tr>
</tbody>
</table>

Note:
Data on schooling from 2010 Population Census, in the same AGEB as data on firms from Economic Census.
But the same $\mathbb{E}(\cdot)$ and K/L

<table>
<thead>
<tr>
<th></th>
<th>Country</th>
<th>Mexico City</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firms</td>
<td>3,643,982</td>
<td>452,102</td>
</tr>
<tr>
<td>Share of total K</td>
<td>100</td>
<td>18</td>
</tr>
<tr>
<td>Share of total L</td>
<td>100</td>
<td>19</td>
</tr>
</tbody>
</table>

Dispersion of TFPR

<table>
<thead>
<tr>
<th></th>
<th>Country</th>
<th>Mexico City</th>
</tr>
</thead>
<tbody>
<tr>
<td>p90 – p10</td>
<td>2.98</td>
<td>3.25</td>
</tr>
<tr>
<td>p75 – p25</td>
<td>1.51</td>
<td>1.57</td>
</tr>
<tr>
<td>dev. std.</td>
<td>1.18</td>
<td>1.25</td>
</tr>
</tbody>
</table>

- “Distortions” in Mexico and Mexico City are similar: in fact, slightly larger in Mexico City.
- This is the expected result, as there are no significant regulatory differences between the two regions.
Q3: given \( E \), will an increase in \( F[l(s,a)] \) raise \( P \)?

- Mexico City (circle) and Mexico have same \( E(.) \); see below. However, years of schooling in Mexico City 33% higher than in Mexico; other indicators of schooling also higher; no data on quality. Note that \( K/L \) in Mexico City and Mexico very similar.

- We compare the physical productivity (TFPQ) of firms of the same type between Mexico City and Mexico

\[
\text{Log } TFPQ_{si} / \text{log average TFPQ in sector s}
\]

<table>
<thead>
<tr>
<th></th>
<th>All firms</th>
<th>Legal-formal</th>
<th>Legal-informal</th>
<th>Legal-semiformal</th>
<th>Semileg-semifor</th>
<th>Illegal-informal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mexico City s.e.</strong></td>
<td>0.141</td>
<td>0.123</td>
<td>0.1888</td>
<td>0.077</td>
<td>0.105</td>
<td>0.079</td>
</tr>
<tr>
<td>s.e.</td>
<td>[0.1069]</td>
<td>[0.0571]</td>
<td>[0.0922]</td>
<td>[0.0767]</td>
<td>[0.0596]</td>
<td>[0.0723]</td>
</tr>
<tr>
<td><strong>Population/km(^2)</strong></td>
<td>0.083</td>
<td>0.079</td>
<td>0.079</td>
<td>0.090</td>
<td>0.085</td>
<td>0/084</td>
</tr>
<tr>
<td>s.e.</td>
<td>[0.019]</td>
<td>[0.0108]</td>
<td>[0.0162]</td>
<td>[0.0149]</td>
<td>[0.0122]</td>
<td>[0.0138]</td>
</tr>
<tr>
<td><strong>Size dummies</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Age dummies</strong></td>
<td>yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>2,231,019</td>
<td>82,148</td>
<td>1,330,573</td>
<td>49,120</td>
<td>143,357</td>
<td>625,821</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.118</td>
<td>0.143</td>
<td>0.043</td>
<td>0.072</td>
<td>0.124</td>
<td>0.053</td>
</tr>
</tbody>
</table>

- Controlling for agglomeration and firm size and age, firms of all types more productive in Mexico City; Average TFPQ difference of 14%, associated with 33% difference in years of schooling.

- Average firm size in Mexico City larger than Mexico (7.4 vs. 4.8 workers). Same result holds for every firm type, e.g., legal-formal (52.1 vs. 32.2), legal-informal (3.3 vs. 2.5) and illegal-informal (4.5 vs. 4.1).

- In parallel, we find that controlling for size and age, within Mexico City, legal and formal firms have higher TFPQs than other firm types (not shown here).
Q4: given $E$, will an increase in $F(s,a)$ decrease informality?

<table>
<thead>
<tr>
<th>Firm type</th>
<th>Mexico</th>
<th>Mexico City (circle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>legal and formal</td>
<td>3.1</td>
<td>3.8</td>
</tr>
<tr>
<td>legal and informal</td>
<td>67.5</td>
<td>63.8</td>
</tr>
<tr>
<td>legal and semi-formal</td>
<td>1.7</td>
<td>1.3</td>
</tr>
<tr>
<td>semi-legal and semi-formal</td>
<td>5.1</td>
<td>5.6</td>
</tr>
<tr>
<td>illegal and informal</td>
<td>22.7</td>
<td>25.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[0-5]</td>
<td>89.7</td>
<td>86.6</td>
</tr>
<tr>
<td>[6-10]</td>
<td>5.8</td>
<td>6.7</td>
</tr>
<tr>
<td>Firm size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[11-50]</td>
<td>3.6</td>
<td>5.1</td>
</tr>
<tr>
<td>[+50]</td>
<td>0.9</td>
<td>1.6</td>
</tr>
<tr>
<td>Dispersion of TFPR (&quot;distortions&quot;)</td>
<td>p90 - p10</td>
<td>2.98</td>
</tr>
<tr>
<td></td>
<td>p75 – p25</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>std dev</td>
<td>1.18</td>
</tr>
</tbody>
</table>

- “Distortions” similar between Mexico and Mexico City (if anything, slightly higher in Mexico City).
- More human capital in Mexico City vs. Mexico does not change the type distribution of firms.
  - share of informal firms in Mexico City (89.3) almost the same as in Mexico (90.2).
  - fewer legal and informal firms in Mexico City but, surprisingly, more illegal and informal firms.
- Issues of transport costs/distance/infrastructure seem to play no role in level of informality.
- Judged by levels of illegality enforcement appears to be, if anything, slightly weaker in Mexico City.
Persistent informality despite more schooling

- In the last 23 years, workers in Mexico have accumulated 2.9 more years of schooling (42%), but labor informality did not change. (Data from household surveys.)
- We have patchy indicators of quality of education (very low relative to OECD, but increasing very slowly).
5. Disentangling “distortions” in $\mathcal{E}(.)$
Discussion

• We measure large wedges in the MRP of K and L across narrowly defined industries (6-digit) and, in case of Mexico City (circle), also narrowly defined geographical space.

• We interpret these wedges as resulting from “distortions” in the environment $\mathbb{E}(.)$.

• We show that these wedges misallocate resources and lower productivity.

• We also show that these wedges induce too many resources into firms with non-salaried and illegal salaried contracts (i.e., informal firms) and/or block resources from being channeled to firms with legal salaried contracts (i.e., formal firms); there are “too many” informal firms and “too few” formal ones.

• (We suspect that there is also “too much” self-employment, but we cannot test this with our data.)

• We show as well large differences in the productivity of different type of informal firms (given size, age and location), implying that there is no one-to-one mapping between informality and productivity.

• Finally, we show that wedges in the MRP of K and L seem to be invariant to stock of human capital (years of schooling), but that, given constant wedges, more human capital increases productivity.
• We conclude that:

✓ low productivity is a result of “distortions” that misallocate resources,
✓ excess informality is also a result of these “distortions”,
✓ more human capital increases productivity but does not reduce informality.

• However, we cannot yet:

✓ establish that the only effect of “distortions” is to change the type distribution of firms,
✓ associate individual policies or regulations that create “distortions” in $f(.)$ with measured wedges in the MRP of K and L because many policies –or market failures– may stand behind observed wedges (no one-to-one mapping).

• **Next step:** to identify whether “distortions” are also inducing resource re-allocation between firms of different sizes, we provide measures of TFPR across firm size and firm type.
Revenue productivity across firm size and firm type

**OLS: log TFPR<sub>i</sub>/log average TFPR<sub>s</sub>; legal-informal [0 – 5] excluded**

(controls for firm age and state; all coefficients significant at 99% level, R-squared = 0.044).

Two broad trends:

- Resources less valuable in informal firms. (Small [0-5] legal informal firms most inefficient of all.)
- Resources more valuable in small and medium size formal and semi-formal firms vs. large ones.
• If the only effect of “distortions” was to change the type distribution of firms in the direction of “too much” informality, we should expect differences in TFPR across firm types, but not across firm sizes.

• However, we also find large differences in TFPRs among firm size.

• In particular, resources are more valuable among medium [11-50] and small [0-5 & 6-10] formal and semi-formal firms vs. large [50+] firms of the same type.

• A first implication is that misallocation and low productivity result from two broad types of “distortions”:
  
  ✓ First, “distortions” that lower productivity by affecting the structure of labor contracts and changing the type distribution of firms in the direction of too much informality.

  ✓ Second, “distortions” that also lower productivity by impeding small and medium formal and semi-formal firms from reaching their optimal size, while channeling too many resources to larger formal and semi-formal firms; but which need not affect informality.
A second implication is that, in the presence of “distortions”, there is no obvious association between firm size and productivity. (In fact, correlation between TFPQ and size is 0.362.)

- on one hand, shifting resources from small informal firms to small formal firms would increase productivity;
- on the other, shifting resources from large to small and medium formal and semi-formal firms would also increase productivity.

A third implication is that the thick left tail of the distribution of firm productivities is mostly accounted for by small informal firms, not by small firms in general, suggesting in turn that:

- many individuals are participating in the labor market as entrepreneurs, but that given their schooling and abilities (i.e., their l(s,a)), these entrepreneurs, and the workers and capital that they use, would be more productive in other firms, and
- in the absence of “distortions”, many of these firms would probably not exist (or would be even smaller).
6. Final observations
• Informality and low productivity jointly determined as outcomes of $\mathbb{E}(.)$; neither chicken nor egg.

• Accumulating human capital increases potential productivity (quality of endowments improves!) and, given constant “distortions”, observed productivity; but need not reduce informality.

• View that “Education will take care of informality” is probably flawed. (Similarly with “Growth will take care of informality”.)

• Critical importance of understanding why $\mathbb{E}(.)$ differs from $\mathbb{E}^*(.)$. Emerging literature looks at individual components or specific regulations of $L$, $C$, $T$, or $Z$ in $\mathbb{E}(.)$, but far from a full understanding of how individual policies interact with each other and translate into observed wedges in MRPs.

• Low productivity and informality over-determined. (Which of the five bullets in the heart killed the guy?)

• That said, in the case of Mexico we broadly find that policies that affect the contractual structure of firms (i.e., that induce too much informality) are quantitatively more relevant than those that affect their size (though some policies do both).

• Discussion very relevant for design and relative importance of reform of:
  ✓ social insurance and labor policies;
  ✓ micro-credits and programs for SMEs;
  ✓ “entrepreneurship” and “formalization” programs;
  ✓ taxation of labor income and special tax regimes for small firms;
  ✓ enforcement of labor and tax regulations.
THANK YOU!