DISCUSSION PAPER Nº IDB-DP-1107

A Framework to Study Agglomeration, Informality, and Firm Dynamics (Discussion Paper)

Juan Pablo Chauvin
Miguel Ángel Talamas Marcos

Inter-American Development Bank
Department of Research and Chief Economist
October 2025

A Framework to Study Agglomeration, Informality, and Firm Dynamics (Discussion Paper)

Juan Pablo Chauvin*
Miguel Ángel Talamas Marcos*

* Inter-American Development Bank

Inter-American Development Bank
Department of Research and Chief Economist

October 2025

Abstract

In this discussion paper we present a framework to analyze how urban agglomeration affects firm informality and the mechanisms behind this relationship. This framework extends a model with extensive and intensive informality margins (Imbert and Ulyssea, 2025) to incorporate agglomeration forces. Firms maximize static profits under Cobb–Douglas technology and choose (i) whether to operate formally or informally, and (ii) among formal operators, the share of workers hired informally. Agglomeration works through two forces: it shifts productivity paths upward and it raises the expected costs of informality (for fully informal operators and for informal hiring within formal firms). These ingredients generate clean cutoff rules for exit and for formalization, and imply that the off-the-books hiring share chosen by formal firms falls with density. In a stationary equilibrium, denser cities display higher average productivity (from level shifts and tighter selection), less informality on both margins, and a greater likelihood of formal entry. Conditional on status, agglomeration acts mainly as a level shifter rather than changing within-firm growth rates.[1]

JEL Classification: E26, J46, O17, R11

Keywords: Agglomeration, Informality, Productivity, Firm Dynamics

[1]We are grateful to Michael Cardona-Rodriguez and Germán Campos for their outstanding research assistance, and to Emilia Bullano for additional research support. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

1 Introduction

Informality remains a defining feature of low- and middle-income economies, shaping tax capacity, worker protection, and the allocation of resources. Our understanding of this phenomenon is quickly evolving: recent research has shown, for instance, that the productivity distributions of formal and informal firms substantially overlap, and that formal firms often hire workers informally (Ulyssea, 2018). Despite this progress, the spatial dimension of informality remains largely understudied. Geographic variation in informality can offer novel insights, particularly given that one of the canonical facts in urban economics is that agglomeration raises productivity, and higher productivity is associated with formality at the firm level. However, the implied connection between density and informality at the aggregate level is a priori ambiguous, because the channels that drive citywide informality—such as entry costs, enforcement, and migration—can offset or reinforce the micro productivity–formality link. This discussion paper presents a framework to study how employment agglomeration relates to firm informality, and how this connection interacts with productivity.

To study this question, we propose a model that embeds the informality-and-dynamics framework of Imbert and Ulyssea (2025) into an urban environment with agglomeration forces. Firms choose whether to operate formally or informally (extensive margin) and, if formal, what share of workers to hire off the books (intensive margin). Agglomeration enters in two places. First, it shifts firms' productivity paths upward through standard agglomeration forces, so establishments in denser cities expect higher productivity today and in the future. Second, it raises the expected costs of informality—both for fully informal operators (via higher detection risk) and for the intensive margin inside formal firms (via steeper expected costs for informal hiring). These features generate simple, threshold-based behavior: a cut-off for exit (low-productivity firms shut down), a cut-off for formalization (informal incumbents formalize once productivity is high enough), and entry composition—both in types and in formal vs. informal status—driven by the productivity shift and stricter enforcement in denser cities.

Building on these ingredients, the model yields sharp, testable predictions for citywide outcomes. Where agglomeration is stronger, average firm productivity is higher because the productivity path shifts up and selection prunes the lower tail, tightening the survival cutoff. Informality falls on both margins: the share of fully informal establishments declines and, among formal firms, the optimal off-the-books hiring share contracts as detection risk steepens. These declines are predicted to be most pronounced for medium and large establishments, for which fixed formalization costs are relatively less binding. On the entry margin, denser cities attract larger and more productive entrants and raise the probability of formal entry. On the exit margin, informal incumbents face higher exit risks in dense places, and conditional exit occurs higher in the size and productivity distributions.

2 Set up

2.1 Environment and Timing

We consider a set of cities j characterized by an agglomeration measure A_j. Time is discrete, $t = 0, 1, 2, \ldots$. Within city j, a continuum of potential establishments (hereafter "firms") may enter and operate. We focus on the partial equilibrium at the city level: output prices p_j and wages w_j are taken as given to focus on the mechanisms of interest. Each incumbent firm has a *status* $s \in \{i, f\}$, where i denotes *informal* (unregistered) and f *formal* (registered). A formal firm can also choose an *intensive* informality share $m \in [0, 1]$ of its employees hired informally.

Within each period, events unfold as follows. First, the establishment learns its current productivity, θ_t, which summarizes how valuable an extra unit of labor would be this period and influences expected profitability going forward. Second, it chooses how to operate. An informal incumbent may pay the fixed cost to become formal, while a formal incumbent selects the share $m \in [0, 1]$ of its workforce hired without a formal contract. Third, taking status and m as given, the firm chooses employment ℓ_t to maximize profits. Fourth, profits are realized. Fifth, the firm compares the value of continuing with the value of shutting down: it exits if its continuation value is negative and, in addition, may be hit by an exogenous "death" shock with status–specific probability. Finally, survivors move to period $t+1$ with productivity updated according to a stochastic process.

2.2 Production, Agglomeration, and Productivity

Technology is Cobb–Douglas, such that output for a firm with productivity θ and labor ℓ is $y = \theta\,\ell^{\alpha}$ with $\alpha \in (0,1)$. Let p_j and w_j denote the city-specific output price and wage. Agglomeration shifts the productivity path according to

$$\ln\theta_t \;=\; \kappa\,g(A_j) \;+\; \rho\ln\theta_{t-1} \;+\; (1-\rho)\ln\nu \;+\; \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0,\sigma^2), \tag{1}$$

where ν is a time-invariant type drawn upon entry from $H(\nu)$, $\rho \in (0,1)$ governs persistence, and $g(A_j)$ is increasing[2]. The parameter κ measures the strength of the influence of agglomeration on productivity levels, such that agglomeration acts as a common shift in log productivity.

2.3 Informality

We follow Ulyssea (2018) and Imbert and Ulyssea (2025) and model informality along two margins—an *extensive* margin (operate as formal or informal) and an *intensive* margin (within formal firms, what fraction of workers are hired off the books). Both margins matter for costs.

On the extensive margin, an informal operator avoids payroll tax costs but faces status–specific fixed and variable costs that reflect limited legal protection, weaker access to formal services, and the risk of being detected. We allow fixed costs and per–worker wedges to depend on the city's agglomeration level A_j, reflecting economies of scale in enforcement that make detection more likely in denser cities. Concretely, informal firms pay a status–specific fixed cost $F_i(A_j)$ and incur an expected per–worker enforcement wedge $\phi_i(A_j) \geq 0$, which rises with A_j.

Formal firms instead pay $F_f(A_j)$ and have to pay statutory payroll contributions $\tau \geq 0$ on their *formal* hires. A formal firm can still choose to hire a fraction $m \in [0,1]$ of its workforce

[2] We assume $g(A) = \ln A$ in the baseline

informally to save on payroll, but doing so raises the chance and cost of detection.The resulting unit labor cost for a formal firm is

$$C_f(m; A_j) \;=\; (1+\tau)(1-m) \;+\; (1-s)\,m \;+\; \kappa_m(A_j)\,m^{1+\eta}, \tag{2}$$

where $s \in (0,1)$ is the private wage saving from an informal contract (relative to a formal one), $\eta > 0$ governs how quickly detection costs ramp up as the informal share grows, and $\kappa'_m(A_j) > 0$ captures that denser cities make intensive informality more costly at the margin. The first term is the wage bill for the formal share $(1-m)$ including payroll τ; the second is the cheaper wage for the informal share m; the third is the expected penalty, rising and convex in m.

For a fully informal firm, the unit labor cost is

$$C_i(A_j) \;=\; 1 - s_i \;+\; \phi_i(A_j), \qquad \phi'_i(A_j) > 0, \tag{3}$$

where $s_i \in (0,1)$ denotes the private saving from paying informal wages outside the formal system, and $\phi_i(A_j)$ is the expected enforcement wedge that grows with A_j.

Fixed operating costs may also differ by status, with the net fixed cost of formality denoted by

$$\Delta F(A_j) \equiv F_f(A_j) - F_i(A_j).$$

Intuitively, $\Delta F(A_j)$ reflects registration fees, bookkeeping, and compliance requirements, while $C_i(\cdot)$ and $C_f(\cdot)$ capture ongoing labor–cost differences and enforcement risk. Higher A_j raises both ϕ_i and κ_m, making informality less attractive both in the extensive and the intensive margin.

3 Single-period Static Problem

Conditional on a city j and a status choice $s \in \{i, f\}$ for the period, a firm with productivity θ chooses its inputs to maximize *current* profits. With Cobb–Douglas technology $y = \theta\,\ell^{\alpha}$ ($\alpha \in (0,1)$), the static problems for informal and formal firms are, respectively

$$\pi_i(\theta; A_j) \;=\; \max_{\ell\geq 0} \Big\{ p_j\,\theta\,\ell^{\alpha} \;-\; w_j\,C_i(A_j)\,\ell \;-\; F_i(A_j) \Big\}, \text{ and} \tag{4}$$

$$\pi_f(\theta; A_j) \;=\; \max_{m\in[0,1]}\;\max_{\ell\geq 0} \Big\{ p_j\,\theta\,\ell^{\alpha} \;-\; w_j\,C_f(m; A_j)\,\ell \;-\; F_f(A_j) \Big\}. \tag{5}$$

Here p_j and w_j are the (city-specific) output price and wage, and the unit labor costs $C_f(m; A_j)$ and $C_i(A_j)$ are defined in equations (2) and (3). Intuitively, the firm trades off the marginal product of labor, $\alpha p\,\theta\,\ell^{\alpha-1}$, against the effective marginal cost $w\,C_s(A_j)$, while bearing the fixed cost $F_s(A_j)$.

For any given labor unit-cost index $C > 0$ (i.e., the wage marked up by taxes or expected penalties), the Cobb–Douglas demand for labor and the associated variable profits are

$$\ell^*(\theta; C) \;=\; \left(\frac{\alpha p\,\theta}{w\,C}\right)^{\frac{1}{1-\alpha}}, \qquad \pi^{\text{var}}(\theta; C) \;=\; (1-\alpha)\,p\,\theta\,\big(\ell^*(\theta; C)\big)^{\alpha}. \tag{6}$$

Two elasticities capture the key trade–off. Holding prices fixed, a 1% increase in productivity raises variable profits by $\frac{1}{1-\alpha}\%$, while a 1% increase in the effective unit labor cost C *reduces* variable profits by $\frac{\alpha}{1-\alpha}\%$. Intuitively, with decreasing returns, more productive firms scale up more than proportionally, but anything that inflates the effective wage (i.e., payroll contributions or enforcement wedges) cuts scale and profits proportionally through C.

Substituting equation (6) into equations (4) and (5) yields a convenient closed form expression:

$$\pi_s(\theta; A_j) \;=\; K(\alpha, p, w)\,\theta^{\frac{1}{1-\alpha}}\,[C_s(A_j)]^{-\frac{\alpha}{1-\alpha}} \;-\; F_s(A_j), \qquad s \in \{i, f\}, \tag{7}$$

where $K(\alpha, p, w) > 0$ is a constant and, for formal firms, $C_f(A_j) \equiv \min_{m\in[0,1]} C_f(m; A_j)$. Thus, conditional on the realized productivity θ, agglomeration A_j enters static profits only through its effects on the effective unit cost $C_s(A_j)$ (via payroll/enforcement wedges) and the fixed cost $F_s(A_j)$.

The intensive margin m for a formal firm is chosen to *minimize* its effective unit cost

$C_f(m; A_j)$ before picking labor. Differentiating equation (2) yields the interior first–order condition

$$-(1+\tau) \;+\; (1-s) \;+\; (1+\eta)\,\kappa_m(A_j)\,m^{\eta} \;=\; 0, \tag{8}$$

with the solution truncated to the admissible set $m \in [0,1]$. Solving for m,

$$\tilde{m}(A) = \left(\frac{\tau + s}{(1+\eta)\,\kappa_m(A)}\right)^{1/\eta}, \qquad m^*(A) = \min\{1, \max\{0, \tilde{m}(A)\}\}. \tag{9}$$

This expression makes the comparative statics transparent. A higher payroll rate τ or a larger private saving s (bigger formal–informal wage gap) pushes m^* up, while stronger visibility/enforcement—captured by a larger $\kappa_m(A_j)$ in denser cities—pushes m^* down. Because $\kappa'_m(A_j) > 0$ by assumption, we obtain $\partial m^*/\partial A_j < 0$: agglomeration reduces the optimal informal share *within* formal firms. Corner solutions are also intuitive: if enforcement is very weak (small κ_m) and payroll is high, $m^* = 1$ (i.e., all hires are off the books); if enforcement is strong enough, $m^* = 0$ (i.e., no informality in the intensive margin). These choices feed back into equation (7) through $C_f(A_j)$ and therefore into the firm's scale $\ell^*(\theta; C_f)$ and current profits.

3.1 Firm Entry

A mass of potential entrants observes a permanent type $\nu \sim H(\nu)$ and an initial draw θ_0 generated by equation (1). Entry options are: do not enter, enter informal, or enter formal. Let entry costs be $C_{e,i}(A) = c_{i0}$ for informal entry and $C_{e,f}(A) = c_{f0} + c_{f1} \ln A$ for formal entry.

Agglomeration A affects potential entrants' choices in two opposing ways. On the upside, higher A lifts expected productivity through the $\kappa g(A)$ term in equation (1), making any operation more profitable. On the downside, higher A also raises the expected costs of operating informally—both the per–worker wedge $\phi_i(A)$ for fully informal firms and the intensive–margin penalty $\kappa_m(A)$ for formal firms that hire some workers off the books. When the *productivity gain* and the *higher informal penalties* dominate any increase in the fixed formal–informal gap $\Delta F(A) \equiv F_f(A) - F_i(A)$, the fraction of entrants that choose to start up *formally* rises with A.

Selection at entry therefore operates along two dimensions. First, because equation (1) shifts the productivity path up in denser markets, the distribution of initial productivity draws θ_0 is stochastically higher where A is larger. This raises the profitability of entry and, conditional on entry, tends to yield entrants with higher realized θ_0. The effect on the composition of permanent types ν is, however, a priori ambiguous. Second, the same rise in enforcement tilts the *status* decision toward formality at the time of entry, so the share entering formally increases with A (under a mild dominance condition stated in Appendix Section A.4). In sum, denser cities feature entrant pools with higher realized productivity at entry and a higher probability of registering formally.

Because A shifts $\ln\theta$ up through $\kappa g(A)$, the distribution of θ_0 in denser cities is stochastically larger. Combined with higher informal penalties, this raises the *relative* attractiveness of formal entry (see equation (18)). The net effect on the composition of entrants by permanent type ν is a priori ambiguous: while the rightward shift in profitability can lower the entry threshold and draw in additional (lower–ν) entrepreneurs, higher enforcement can tilt status toward formality. Under the dominance condition (18), the *share* entering formally rises with A.

3.2 Firm Growth, Formalization, and Exit

Let $V_s(\theta; A_j)$ be the value of an incumbent with status $s \in \{i, f\}$ and current productivity θ in city j. After earning static profits $\pi_s(\theta; A_j)$ from equation (7), the firm decides whether to (i) continue next period under its current status, (ii) switch status (an informal firm may formalize by paying the net fixed gap $\Delta F(A_j) \equiv F_f(A_j) - F_i(A_j)$ and a formal firm may revert to informal without that payment in the baseline), or (iii) shut down. Continuation is subject to a status–specific exogenous death shock $\delta_s \in [0, 1)$, and future productivity follows the agglomeration–shifted process in equation (1). With discount factor $\beta \in (0, 1)$, the Bellman equations are

$$
\begin{aligned}
V_i(\theta;A) = \max\Big\{ &\underbrace{\pi_i(\theta;A) + \beta(1-\delta_i)\,\mathbb{E}[V_i(\theta';A)\mid\theta]}_{\text{remain informal}}, \\
&\underbrace{\pi_f(\theta;A) - \Delta F(A) + \beta(1-\delta_f)\,\mathbb{E}[V_f(\theta';A)\mid\theta]}_{\text{formalize now}},\ 0\Big\}.
\end{aligned}
\tag{10}
$$

$$
\begin{aligned}
V_f(\theta;A) = \max\Big\{ &\underbrace{\pi_f(\theta;A) + \beta(1-\delta_f)\,\mathbb{E}[V_f(\theta';A)\mid\theta]}_{\text{remain formal}}, \\
&\underbrace{\pi_i(\theta;A) + \beta(1-\delta_i)\,\mathbb{E}[V_i(\theta';A)\mid\theta]}_{\text{revert to informal}},\ 0\Big\}.
\end{aligned}
\tag{11}
$$

Each expression within the maximization represents the value of continuing under that particular choice. Remaining informal avoids payroll but keeps the enforcement wedge; formalizing pays the fixed gap $\Delta F(A)$ and then benefits from lower enforcement risk and access to formal inputs; shutting down yields zero. Exogenous death (δ_s) captures events such as landlord decisions or non–modeled shocks that remove the firm regardless of profitability. Crucially, the agglomeration index A raises expected future productivity through $\kappa g(A)$ in equation (1) and increases the expected costs of informality through $\phi_i(A)$ and $\kappa_m(A)$, so A affects both the static payoff π_s and the option value of continuing.

Because current profits (equation (7)) are strictly increasing in θ and the productivity process in (1) is monotone (higher θ today stochastically implies higher θ' tomorrow), the problem exhibits single–crossing. As a result, optimal behavior is governed by two cutoffs. First, there is an *economic exit* threshold $\theta^*(A)$ such that the firm exits if and only if $\theta < \theta^*(A)$. A higher cutoff means tougher selection. Second, there is also a *formalization* threshold $\bar{\theta}(A)$ such that an informal incumbent formalizes if and only if $\theta \geq \bar{\theta}(A)$.

Putting these pieces together, agglomeration simultaneously (i) shifts firms' productivity paths to the right and (ii) trims the lower tail through tougher selection, while (iii) making informality less attractive at both margins. In stationary equilibrium, denser cities therefore feature entrants and survivors that are more productive and more likely to operate formally, higher pre–exit productivity among exiters, and lower intensive informality within formal firms.

4 Equilibrium and comparative statics

4.1 Stationary Equilibrium

Fix (A_j, p_j, w_j). A stationary city equilibrium is one in which, period after period, firms follow the same decision rules and the composition of firms by productivity and status no longer changes. Concretely, there exist policy functions $m^*(A_j)$, $\bar{\theta}(A_j)$, and $\theta^*(A_j)$ solving the static cost problem in equation (2) and the dynamic problems in equations (10) and (11). Under these policies, the joint distribution of (θ, s) is invariant given the productivity law of motion (equation (1)), optimal entry, and endogenous formalization and exit. The implied city aggregates—average productivity, extensive and intensive informality, and the rates of entry, growth, and exit—are therefore constant over time. Because prices and wages are taken as given, agglomeration shapes the stationary allocation only through its effects on the productivity process, on effective unit labor costs, and on status–specific fixed operating costs that enter the firms' optimal rules.

4.2 Comparative Statics and Implied Reduced-Form Relationships

Agglomeration shapes citywide results through three main channels. First, it lifts firms' productivity paths. In equation (1), the term $\kappa g(A_j)$ shifts $\ln \theta_t$ upward in denser cities, so a given firm expects higher productivity today and tomorrow. Holding costs fixed, the profit formula (equation (7)) then implies a larger optimal scale $\ell^*(\theta; C)$ and higher variable profits, which shows up across firms as a rightward shift of the TFPR distribution and higher citywide averages.

Second, agglomeration makes informality more expensive. On the extensive margin, the expected per–worker wedge $\phi_i(A_j)$ in equation (3) rises with A_j. On the intensive margin within formal firms, the penalty term $\kappa_m(A_j)\, m^{1+\eta}$ in equation (2) grows with A_j, so the cost–minimizing informal share $m^*(A_j)$ in equation (9) falls. Together these forces lower the formalization threshold $\bar{\theta}(A_j)$ in the dynamic problem (equations (10) and (11)) and reduce $m^*(A_j)$, yielding less informality at both margins in denser cities.

Third, denser markets can tighten selection by raising the bar to survive: with higher fixed costs and stricter enforcement, low–θ firms find it harder to stay in the market. Concretely, when (in the neighborhood of the exit margin) fixed operating costs and enforcement pressures are at least as high—and plausibly higher—in denser places, the static break–even in equation (7) shifts up with A_j. Firms must reach a higher productivity to stay afloat, so the continuation value turns negative at a larger θ and the exit cutoff $\theta^*(A_j)$ rises. This "pruning" of low–θ firms raises the average productivity of survivors and, among exiting firms, implies higher pre–exit productivity. In short, equation (1) shifts productivity up, equations (2) and (3) raise the cost of informality, and—under the condition just described—equation (7) tightens the profitability floor, together delivering higher productivity and lower informality where agglomeration is stronger.

These mechanisms translate into clear reduced–form predictions that we take to the data in our empirical section. Where agglomeration is stronger (higher A_j), the upward shift in productivity paths implied by equation (1) and the profit scaling in equation (7) should raise average productivity across firms. At the same time, because agglomeration makes informality costlier at both margins, denser cities should display a smaller share of informal establishments and, among formal firms, a lower informal hiring share. These declines are expected to be most pronounced for larger establishments, for which fixed formalization costs are relatively less binding.

On the entry margin, the same forces predict positively selected entrant pools—bigger and more productive firms, and a higher probability of entering formally. Among survivors, we expect limited within–firm TFPR acceleration once status is held fixed, since equation (1) primarily acts as a level shifter rather than altering growth rates.

Finally, the model links exit patterns to a simple cutoff rule. A firm exits when its continuation value turns nonpositive. Let $W_s(\theta; A)$ denote the best continuation value if the firm operates with status s this period (the maximum over the two nonexit branches in equations (10) and (11)). The exit threshold $\theta^*(A)$ is the smallest θ such that $W_s(\theta; A) \geq 0$.

To build intuition, consider the static break-even first. If the firm ignored future options, it would exit whenever current profits are negative. Let $\theta_s^\dagger(A)$ solve $\pi_s(\theta_s^\dagger(A); A) = 0$. This is the static break-even productivity for status s. In the dynamic problem, the true cutoff $\theta^*(A)$ lies near this level but adjusts for the option value of continuing—the chance that θ' improves. Tighter operating requirements or stronger enforcement raise the static break-even through higher $F_s(A)$ or a larger informal wedge $\phi_i(A)$. The dynamic cutoff moves with it, implying tougher selection in denser cities.

5 Conclusion

This paper presents a dynamic framework that links agglomeration to firm informality through two channels: a level shift in productivity paths and higher expected costs of operating or hiring informally. These features generate threshold policies for survival and formalization, and an intensive choice for informal hiring inside formal firms that declines with density. In the stationary allocation, denser cities combine a rightward shift of the productivity distribution with tighter pruning of the lower tail, while informality falls on both margins and formal entry becomes more likely.

Beyond these aggregate patterns, the framework delivers concrete organizing principles for measurement and testing. First, conditional on status, agglomeration operates as a level shifter rather than altering within–firm growth rates. Second, the model clarifies how entry, formalization, and exit interact with enforcement to shape the joint distribution of productivity and informality. These implications motivate reduced–form exercises that separate level shifts from selection and that contrast intensive vs. extensive margins across the size distribution.

References

Imbert, C. and Ulyssea, G. (2025). Rural Migrants and Urban Informality: Evidence from Brazil. *Econometrica*.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108(8):2015–2047.

A Model Appendix

A.1 Solution Details

A.1.1 Static profits, optimal scale, and the intensive margin

Fix a city j and suppress j in notation where no confusion arises. With Cobb–Douglas technology $y = \theta\,\ell^{\alpha}$ and unit labor cost index $C > 0$, the static problem yields the labor FOC $\alpha p\theta\ell^{\alpha-1} = wC$. Hence

$$\ell^{*}(\theta; C) = \left(\frac{\alpha p\theta}{wC}\right)^{\frac{1}{1-\alpha}}, \qquad \pi^{\text{var}}(\theta; C) = (1-\alpha)\,p\,\theta\,\big(\ell^{*}(\theta; C)\big)^{\alpha}. \tag{12}$$

Substituting (12) gives the closed form

$$\pi^{\text{var}}(\theta; C) = K(\alpha, p, w)\,\theta^{\frac{1}{1-\alpha}}\,C^{-\frac{\alpha}{1-\alpha}}, \quad K(\alpha, p, w) \equiv (1-\alpha)\,\alpha^{\frac{\alpha}{1-\alpha}}\,p^{\frac{1}{1-\alpha}}\,w^{-\frac{\alpha}{1-\alpha}}. \tag{13}$$

Plugging into the status-specific problems in the main text, equations (4)–(5), yields the compact profit expression in equation (7):

$$\pi_s(\theta; A) = K(\alpha, p, w)\,\theta^{\frac{1}{1-\alpha}}\,[C_s(A)]^{-\frac{\alpha}{1-\alpha}} - F_s(A), \qquad s \in \{i, f\}. \tag{14}$$

On the intensive margin inside formal firms, equation (2) is

$$C_f(m; A) = (1+\tau)(1-m) + (1-s)m + \kappa_m(A)\,m^{1+\eta}.$$

Differentiating gives the interior FOC in (8):

$$-(1+\tau) + (1-s) + (1+\eta)\kappa_m(A)m^{\eta} = 0 \iff (1+\eta)\kappa_m(A)m^{\eta} = \tau + s,$$

so the unconstrained optimizer is

$$\tilde{m}(A) = \left(\frac{\tau + s}{(1+\eta)\kappa_m(A)}\right)^{1/\eta}, \qquad m^{*}(A) = \min\{1, \max\{0, \tilde{m}(A)\}\}. \tag{15}$$

Because $\kappa_m'(A) > 0$, $m^*(A)$ is (weakly) decreasing in A, and strictly decreasing away from corners. The cost index entering (7) is $C_f(A) \equiv C_f(m^*(A); A)$.

A.1.2 Productivity and stationarity

The productivity process in the main text (equation (1)) is

$$\ln \theta_t = \kappa\, g(A) + \rho\, \ln \theta_{t-1} + (1-\rho) \ln \nu + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

with $\rho \in (0,1)$. Conditional on (A, ν), $\ln \theta_t$ is AR(1) and admits a unique stationary distribution, Gaussian with mean $\mu(A,\nu) = \frac{\kappa\, g(A) + (1-\rho)\ln \nu}{1-\rho}$ and variance $\sigma^2/(1-\rho^2)$. Hence θ_t is lognormal in stationarity and for any finite $\gamma > 0$,

$$\mathbb{E}\big[\theta_t^\gamma \mid A, \nu\big] = \exp\Big(\gamma \mu(A,\nu) + \tfrac{1}{2}\gamma^2 \tfrac{\sigma^2}{1-\rho^2}\Big) \;<\; \infty. \tag{16}$$

In particular, the moment $\gamma = \frac{1}{1-\alpha}$ required by (13) exists, ensuring finite expected static profits and value functions under $\beta \in (0,1)$.

A.2 Dynamic Program, Existence, and Threshold Policies

A.2.1 Timing and Bellman equations

Within a period, status (and m for $s = f$) is chosen *before* production (main text, "Environment and Timing"). Hence switching status affects current profits. Let $V_s(\theta; A)$ denote the value of an incumbent with current status $s \in \{i, f\}$ and productivity θ in a city with index A. Let δ_s be the exogenous death hazard and $\beta \in (0,1)$ the discount factor. The Bellman equations are given by equations 11 and 10 in the main text.

A.2.2 Existence and monotonicity

Define the operator T mapping (V_i, V_f) to the right-hand sides of (10)–(11). Under (16), there exist $a, b > 0$ and $\gamma = \frac{1}{1-\alpha}$ such that

$$\mathbb{E}\big[\theta'^{\gamma} \mid \theta\big] \;\leq\; a + b\,\theta^{\gamma\rho}.$$

Hence the Bellman operator is a contraction on the space of functions with at most linear growth in θ^{γ} under a weighted sup norm (with discount factor $\beta \in (0,1)$ and survival rates $1 - \delta_s$), delivering a unique bounded solution (V_i, V_f).

Both $\pi_s(\theta; A)$ and the continuation terms are strictly increasing in θ. With the Markov kernel from (1) satisfying monotone likelihood ratio in θ, $\theta \mapsto \mathbb{E}[V_s(\theta'; A) \mid \theta]$ is increasing. Therefore $V_s(\theta; A)$ is strictly increasing in θ for $s \in \{i, f\}$.

A.2.3 Threshold structure

Define the net formalization surplus $\Delta V(\theta; A) \equiv V_f(\theta; A) - V_i(\theta; A)$. As $\pi_f(\theta; A) - \pi_i(\theta; A)$ is strictly increasing in θ (from (7)) and continuation values are increasing in θ, $\Delta V(\theta; A)$ is strictly increasing. There exists a unique formalization threshold $\bar{\theta}(A)$ such that an informal incumbent formalizes iff $\theta \geq \bar{\theta}(A)$.

Let $W_s(\theta; A)$ be the best continuation value if the firm operates with status s this period (i.e., $W_s(\theta; A) = \max\{\text{first two terms in } V_s\}$). Because $W_s(\theta; A)$ is strictly increasing in θ and $W_s(0; A) \leq 0$, there is an exit cutoff $\theta^*(A)$ (possibly status-dependent) such that the firm exits iff $\theta < \theta^*(A)$. With the timing here, $\theta^*(A)$ is the smallest θ for which $W_s(\theta; A) \geq 0$.

Intuition. Productivity shapes the whole calculus: higher θ raises current profits, improves the option value of continuing (because future θ' stochastically increases), and thus makes both survival and formalization more attractive. The max operator turns these monotone differences into clean cutoff rules.

A.3 Selection Tightness and the Economic Cutoff

Ignoring dynamics, the static break-even θ for status s solves $K\,\theta^{\frac{1}{1-\alpha}} C_s(A)^{-\frac{\alpha}{1-\alpha}} = F_s(A)$. Thus

$$\theta_s^\dagger(A) = \left(\tfrac{F_s(A)}{K(\alpha,p,w)}\right)^{1-\alpha} [C_s(A)]^\alpha. \tag{17}$$

If $F_s'(A) \geq 0$ or $C_s'(A) \geq 0$, then $\partial\theta_s^\dagger/\partial A \geq 0$. In the dynamic problem, $\theta^*(A)$ tracks $\theta_s^\dagger(A)$ up to continuation terms: a higher fixed burden or higher effective unit costs raise the profitability floor and prune the lower tail.

A.4 Entry, Sorting, and the Status of Entrants

A potential entrant draws type ν from $H(\nu)$, draws θ_0 from (1), and then chooses status or not to enter, paying $C_{e,i}(A) = c_{i0}$ if informal and $C_{e,f}(A) = c_{f0} + c_{f1}\ln A$ if formal. Let $J_s(\nu, \theta_0; A)$ be the value of entering with status s, net of entry costs. Because A shifts $\ln\theta_t$ by $\kappa g(A)$ and raises informal penalties through $\phi_i(A)$ and $\kappa_m(A)$, J_f increases in A both directly (higher θ levels) and relatively (informality less attractive). Under the mild dominance condition

$$\frac{\partial}{\partial A}\Big[\pi_f(\theta; A) - \pi_i(\theta; A)\Big] \;\geq\; c_{f1}/A \quad \text{for } \theta \text{ in the relevant support,} \tag{18}$$

the probability of *entering formally* rises with A. Because θ_0 is stochastically larger in higher A (level shift in (1)), entrants in denser places are positively selected on (ν, θ_0) and more likely to register.

Intuition. Agglomeration (i) raises the entire productivity path and (ii) tightens informal enforcement. Both forces tilt the status calculus at the door towards formality. The relative strength of these forces determines whether the larger fixed gap $\Delta F(A)$ can overturn the tendency; (18) rules that out.

A.5 Comparative Statics in A

A.5.1 Intensive margin $m^*(A)$ and relative costs

Equation (15) implies $\partial m^*/\partial A < 0$ away from corners, and thus $C_f(A)$ rises with A via two channels: a larger penalty $\kappa_m(A)$ and a smaller $m^*(A)$ that reduces the payroll saving. For fully informal firms, $C_i'(A) = \phi_i'(A) > 0$. Therefore

$$\frac{\partial}{\partial A} \log \frac{C_f(A)}{C_i(A)} = \frac{C_f'(A)}{C_f(A)} - \frac{\phi_i'(A)}{C_i(A)} \quad \text{has ambiguous sign a priori.} \tag{19}$$

We impose the identifying assumption used in the main text:

Assumption A1 (Enforcement bites more on the extensive margin). For A in the empirically relevant range,

$$\frac{\partial}{\partial A}\Big(\pi_f(\theta; A) - \pi_i(\theta; A)\Big) \;\geq\; 0 \qquad \text{for all } \theta \text{ in the support.} \tag{20}$$

A sufficient condition is $\phi_i'(A)$ large enough relative to the increase in $C_f(A)$, or, more directly, $\frac{\partial}{\partial A} \log \frac{C_f(A)}{C_i(A)} \leq 0$.

A.5.2 Formalization threshold $\bar{\theta}(A)$

Let $\Delta V(\theta; A) = V_f(\theta; A) - V_i(\theta; A)$. Using (10)–(11),

$$\Delta V(\theta; A) = \big[\pi_f(\theta; A) - \pi_i(\theta; A)\big] - \Delta F(A) + \beta\Big((1-\delta_f)\mathbb{E}[V_f(\theta'; A) \mid \theta] - (1-\delta_i)\mathbb{E}[V_i(\theta'; A) \mid \theta]\Big).$$

Under Assumption A1 and with the agglomeration shift $\kappa g'(A) > 0$ in (1), the bracketed static difference is weakly increasing in A, and continuation values rise with A because the entire θ path shifts up. Hence $\partial_A \Delta V(\theta; A) \geq 0$, and the unique root $\bar{\theta}(A)$ solving $\Delta V(\bar{\theta}(A); A) = 0$ is (weakly) decreasing in A:

$$\frac{\partial \bar{\theta}}{\partial A} \;\leq\; 0. \tag{21}$$

If $\Delta F'(A) \leq 0$ (registration made easier in dense places), the inequality is strict.

A.5.3 Exit cutoff $\theta^*(A)$

Let $\theta_s^\dagger(A)$ be the static break-even in (17). If $F_s'(A) \geq 0$ or $C_s'(A) \geq 0$, then $\theta_s^\dagger(A)$ rises with A. Because A also shifts θ up via $\kappa g(A)$, the dynamic $\theta^*(A)$ increases with A whenever static tightening dominates the level shift near the lower tail. Sufficient conditions are $F_s'(A) \geq 0$ or $\phi_i'(A) > 0$ large enough. Therefore,

$$\frac{\partial \theta^*}{\partial A} \;\geq\; 0. \tag{22}$$

A.5.4 Within-firm TFPR growth conditional on status

From equation (1),

$$\ln \theta_{t+1} = \kappa g(A) + \rho \ln \theta_t + (1-\rho)\ln \nu + \varepsilon_{t+1},$$

so the city-specific stationary mean is $\mu_A = \frac{\kappa g(A) + (1-\rho)\ln \nu}{1-\rho}$. Therefore,

$$\mathbb{E}\big[\ln \theta_{t+1} - \ln \theta_t \,\big|\, \ln \theta_t, A\big] = \kappa g(A) - (1-\rho)\big(\ln \theta_t - \mu_A\big).$$

Agglomeration A shifts the level μ_A but does not alter the persistence ρ. Hence, holding status (and thus C_s) fixed, there is no differential *slope* of expected TFPR growth by A: cross-city differences in average growth at a point in time reflect distance to the city-specific mean μ_A, not a different dynamic coefficient.

A.6 Distributional Accounting and a Quantile Mapping

Let F_A denote the stationary CDF of $\ln \mathrm{TFPR}$ in cities with agglomeration A. Two forces connect F_{A_H} to F_{A_L}: an additive location shift $\Delta_A \equiv \kappa\big[g(A_H) - g(A_L)\big]$ in $\ln \theta$, and a left-tail truncation difference due to $\theta^*(A_H) \geq \theta^*(A_L)$. Under the maintained assumption that entry and ν do not change the *shape* of the latent productivity distribution across A (only its location), and that selection operates as a lower-tail truncation, the survivor distributions satisfy

$$Q_{A_H}(u) = Q_{A_L}\big(S + (1-S)u\big) + \Delta_A, \qquad u \in (0,1), \tag{23}$$

where Q_A is the quantile function for $\ln \mathrm{TFPR}$ among survivors at A, and $S \in [0,1)$ equals the increment in the survivor share cut off at the bottom when moving from A_L to A_H. Equation (23) is exact under common-shape log-location families and sharp selection; it is a useful approximation otherwise, and is what we implement in the descriptive decomposition. In the baseline, we set dilation to one; if dispersion responds to A, a multiplicative dilation term can be activated.

A.7 Propositions and Proofs

Proposition 1 (Value existence, monotonicity, and thresholds).

Under (1) with $\rho \in (0,1)$, $\beta \in (0,1)$, and finite moment (16), the Bellman operator defined by (10)–(11) admits a unique solution (V_i, V_f). For each fixed A, $V_s(\cdot; A)$ is strictly increasing in θ, and optimal policies are characterized by two cutoffs: an exit threshold $\theta^*(A)$ and a formalization threshold $\bar{\theta}(A)$, both unique.

Proof. Contraction follows from Blackwell: monotonicity is immediate; discounting with $\beta < 1$ and bounded growth from (16) give a sup-norm contraction on the space of functions with linear growth in $\theta^{1/(1-\alpha)}$. Increasing differences in (θ, choice) imply single-crossing; existence and uniqueness of thresholds follow from the strict monotonicity of V_s and the max operator. □

Proposition 2 (Intensive margin and comparative statics).

For $\eta > 0$ and $\kappa_m'(A) > 0$, the optimal intensive share $m^*(A)$ in (15) is weakly decreasing in A and strictly decreasing away from corners. Consequently, $C_f(A)$ is weakly increasing in A.

Proof. Follows by implicit function theorem on $(1+\eta)\kappa_m(A)m^\eta = \tau + s$ in the interior, and by monotonicity of the objective at corners. *Corner behavior.* When $m^*(A) = 0$, $C_f(A) = (1+\tau)$ locally (flat in A); when $m^*(A) = 1$, $C_f(A) = 1 - \sigma + \kappa_m(A)$ and is strictly increasing in A. □

Proposition 3 (Formalization and exit comparative statics).

If Assumption A1 holds and either $\Delta F'(A) \leq 0$ or the agglomeration shift $\kappa g'(A)$ is sufficiently large, then $\bar{\theta}'(A) \leq 0$. If, in addition, $F_s'(A) \geq 0$ or $\phi_i'(A) > 0$, then $\theta^{*\prime}(A) \geq 0$.

Proof. Differentiate $\Delta V(\bar{\theta}(A); A) = 0$ using the envelope theorem. Under A1 and $\Delta F'(A) \leq 0$ (or a sufficiently large level shift in θ), $\partial_A \Delta V \geq 0$ while $\partial_\theta \Delta V > 0$ by single-crossing, so $\bar{\theta}'(A) \leq 0$. For exit, differentiate $W_s(\theta^*(A); A) = 0$. If $F_s'(A) \geq 0$ or $C_s'(A) \geq 0$ near θ^*, then $\partial_A W_s \leq 0$ while $\partial_\theta W_s > 0$, implying $\theta^{*\prime}(A) \geq 0$. □

Proposition 4 (Within-firm growth conditional on status).

Holding status s fixed, expected TFPR growth does not vary with A; agglomeration acts as a level shifter.

Proof. Take expectations in (1) conditional on (A, s) and subtract $\ln \theta_{t-1}$; the A term is time-invariant. □