DISCUSSION PAPER SERIES

IZA DP No. 3550

# Formality, Informality, and Social Welfare

John Bennett

June 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

# Formality, Informality, and Social Welfare

**John Bennett**
*Brunel University*
*and IZA*

Discussion Paper No. 3550
June 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3550
June 2008

# ABSTRACT

## Formality, Informality, and Social Welfare*

An industry is modeled in which entrepreneurs, who are heterogeneous in ability, may produce formally or informally. It is shown how the formal-informal mix depends on the distribution of ability, product demand and various parameter values. The industry equilibrium is compared to one in which informality is prohibited. With relatively high product demand, the effect of entrepreneurs being free to choose informality is that consumer surplus and total employment are reduced, but profit is redistributed towards more able entrepreneurs. With relatively low product demand the opposite effects obtain. We also show that informality may be a built-in stabilizer or destabilizer.

JEL Classification: O17, D2

Keywords: formality, informality, social welfare

Corresponding author:

John Bennett
Department of Economics and Finance
and Centre for Economic Development and Institutions (CEDI)
Brunel University
Uxbridge, UB8 3PH
United Kingdom
E-mail: john.bennett@brunel.ac.uk

* I am grateful for helpful comments by Saul Estrin and Stephanie Levy, and by participants at the Third IZA/World Bank Conference on 'Employment and Development,' Skhirat, Morocco, May, 2008. I also thank the Leverhulme Trust for financial support.

# Formality, Informality, and Social Welfare

## 1 Introduction

The literature on informality in developing economies has to a large extent focused on its employment aspect. This reflects the critical importance of informal employment, which comprises about 60% of the labour force in these economies. Though proportionately smaller than informal employment, comprising nearly 40% of GDP, the output from informal activity also plays a fundamental role, perhaps meriting more attention from researchers than it has so far received.[1] In this paper we examine how the pattern of informality/formality in an industry is related to the distribution of entrepreneurial abilities and the strength of the demand for the industry's output, as well as to various parameters relating to the relative costs and benefits affecting the entrepreneur's choice between formality and informality. We use this framework to consider welfare aspects of the industry equilibrium, an issue that seems to have been neglected in the literature.

Our analysis relates only to activities in which formality-versus-informality is an active choice for an entrepreneur. It is not meant to apply to the lower tier of informal activity, which has low value added, often only involving resale, with low capital-intensity and almost no paid employment. Rather, it relates to informal activity that can make profit, is more capital-intensive, involves paid employment and may entail competition with formal firms (see Fields, 2005, on this distinction). This tier of informal activity can be seen as the equivalent of the

[1] These figures, which are widely cited, are taken from Schneider and Enste (2000). A more recent figure given by Schneider (2006) is that the 'shadow economy,' defined as market-based legal production, constitutes 38.7% of GDP, on average, for 96 developing economies.

small-scale entrepreneurial sector in developed economies (Maloney, 2004).

We characterize a formal firm as one that provides social benefits for workers (or, equivalently, pays a minimum wage rate that does not apply to informal firms), as well as having to pay a fee for firm registration. Thus, in the terminology of Bourgignon (2005), formal firms provide 'good' jobs, while informal firms provide 'bad' ones. We assume, however, that a formal firm obtains a productivity benefit from greater access to public goods. In practice, formality is also strongly associated with larger size, which is often treated as its defining characteristic.[2] Formal regulations may only apply to firms above a specified size, as in India (see Ahsan and Pages, 2007), and, more generally, an entrepreneur operating informally may be inhibited from expanding the firm by the fear of attracting the attention of the authorities (Fortin, Marceau and Savard,1997; World Bank, 2007). To capture this size factor we make the specific assumption that the firm operated by any given entrepreneur would use twice as much capital and labour with formality as it would with informality. Nonetheless, the amount of labour an entrepreneur employs per unit of capital is assumed to depend on a parameter reflecting his or her entrepreneurial ability. Given the distribution of the ability parameter across the population of potential entrepreneurs, we determine the industry equilibrium in terms of which entrepreneurs enter and what status, formal or informal, is chosen by each.

We examine how the formal and informal components of output supply depend

[2]There is no generally agreed way to define informality, though lack of registration is included in most definitions. The two approaches most commonly taken are also to specify that (a) workers are not covered by social security benefits or (b) the firm does not exceed some employment size threshold. For example, definition (a) has been used by the South African government and (b) by the Moroccan one (Walther and Filipiak, 2007). Although we use (a) for most of our analysis, in the penultimate section we consider the effects of instead using (b).

on the distribution of entrepreneurial ability, the output price and the various parameters of the model. When deciding whether to enter formally, an entrepreneur does not internalize either the social benefits his or her firm provides to workers or the registration fee that is a transfer to the state. (If they were internalized, formality would always be chosen.) We distinguish four cases, according to whether, for a formal firm, the cost of social benefit provision and the cost of registration is high or low (each measured relative to a critical value). We then focus on a baseline case - a high social benefit cost but low registration cost - for combinations of parameter values that result in a mix of formality and informality.

We compare the industry equilibrium when each (producing) entrepreneur is free to choose either formality or informality, with one in which informality is prohibited. Having informality as an option is found to cause some entrepreneurs to choose informality who would otherwise have chosen formality, and some to produce informally who would otherwise not have produced at all. We assume that for a sufficiently high output price there is a binding constraint on the supply of entrepreneurs, either because the pool of entrepreneurs is itself relatively limited or because there is an aggregate constraint on a complementary factor (we develop the example of a capital constraint). It is then found that informal output supply may be non-monotonic in price. At relatively low prices there is only formal supply; at intermediate prices there is both formal and informal supply; while at high prices there is only formal supply.

The impact of informality on total output supply can go either way: for a relatively low range of product demand it causes total output to rise, but for a relatively high range of product demand it causes total output to fall. A similar result is holds for the effect on total employment. Moreover, informality can be a

built-in stabilizer or destabilizer in the face of fluctuating product demand: there is an intermediate range of demand for which it is a built-in stabilizer, but outside this range it is a built-in destabilizer.

We also consider the welfare effects of informality being an option for entrepreneurs. The effect on consumer surplus can be of either sign - negative when product demand is high, but positive if product demand is low. The effect on the labour market is that the number of good jobs is reduced, and this is not necessarily counteracted by an increase in the number of bad jobs: total employment rises if product demand is low, but falls if product demand is high. Aggregate profit may be affected in either direction, but if demand is in the higher range informality causes a redistribution of profit towards more efficient entrepreneurs, whereas if it is in the lower range the redistribution is away from the more efficient entrepreneurs, which we suggest may be damaging for investment and growth.[3]

Thus, with high product demand the effect of informality being a choice option is to reduce consumer surplus, total employment and the number of good jobs, but to redistribute profit towards entrepreneurs who may be able to use it more effectively for investment. With low product demand the effect of informality being an option is to raise consumer surplus and employment, but to reduce the number of good jobs and to redistribute profit away from the more able entrepreneurs.

In the literature analyzing informality, heterogeneity of entrepreneurial ability was introduced by Rauch (1991), using a variation of the Lucas (1978) model in which ability is represented by a multiplicative parameter on output. He thus ex-

[3]Informality may also, in principle, have a dynamic benefit as a possible stepping-stone to formality; that is it may allow an entrepreneur to test the market and his or her own ability without incurring sunk costs. Without such a stepping stone the entrepreneur may not enter. This argument is analyzed by Bennett and Estrin (2007).

plains why a 'missing middle' in the size distribution of firms obtains. Our formulation shares with Rauch the assumption of a minimum wage rate (or, equivalently, social benefit provision) only in the formal sector, though we characterize entrepreneurial ability differently. However, at the cost of suppressing the endogenous choice of size for each firm that Rauch analyzes, we are able to examine in detail the product market and welfare ramifications.

Various authors have analysed informality using the Lucas-Rauch formulation of entrepreneurial ability. These include Fortin, Marceau and Savard (1997), who generate a gap in the size distribution of firms as a result of expenditure by informal firms on avoiding detection, and Amaral and Quintin (2006), who endogenize skills by formulating an overlapping-generations model with occupational choice. Other building blocks of our model are also common in the literature, including higher wage rates paid by formal than by informal firms (e.g., Goldberg and Pavcnik, 2003; Banerji and Jain, 2007) and the productivity benefit of formality (e.g., Amaral and Quintin, 2006; Loayza and Rigolini, 2006). Factors that we do not include but which are also found in the literature include taxation (de Paula and Scheinkman, 2006), differential capital costs (Straub, 2005), bribery (Ihrig and Moe, 2004) and job matching (Bosch, 2007). One paper that, like ours, focuses on the product market is that of Banerji and Jain (2007). However, their focus is different, examining the endogenous differences in the quality of output between informal and formal firms.

In Section 2 the model is set up. Section 3 considers the implications of the model for the formal-informal mix, while Section 4 examines welfare issues. Section 5 discusses the robustness of the analysis to alternative assumptions relating to the definition of informality, the relative sizes of parameters, the way social benefits are

paid for, and the characterization of entrepreneurial ability. Section 6 concludes and an appendix contains proofs.

# 2 The Model

Consider a competitive industry, with free entry, producing a homogeneous good. A pool of entrepreneurs exists for this industry, each of whom may choose to run a firm with informal status, or to run a firm with informal status, or to stay out of production. Running one firm requires an entrepreneur's full attention. A firm is a price-taker in the product market and in input markets. To allow for heterogeneity of entrepreneurial abilities, assume that a cost coefficient $\lambda$ is attached to each entrepreneur, where $\lambda > 0$, and that $\lambda$ varies across the pool of entrepreneurs. A higher value of $\lambda$ reflects lower entrepreneurial ability (we refer to an entrepreneur with a given $\lambda$ as a '$\lambda$-entrepreneur').[4]

If a firm is informal it employs $k$ units of capital and produces one unit of output. Setting the capital cost at unity, its capital cost is therefore $k$. In conjunction with the $k$ units of capital, a $\lambda$-entrepreneur employs $\lambda$ units of labour to produce the unit of output, thus incurring a wage cost of $\lambda w$, where $w$ is the wage rate. Writing $p$ for the unit price of output, a $\lambda$-entrepreneur who chooses informality earns profit

$$\pi^I = p - w\lambda - k. \tag{1}$$

Instead, any entrepreneur may choose formal status for a firm. For a $\lambda$-

[4]Thus, in contrast to Lucas (1978) and Rauch (1991), we model differences in ability through a costs, rather than a multiplicative parameter on output. Diiferences in entrepreneurual ability are characterized in cost terms by Jovanovic (1982) and by the literature that his contribution generated (see Parker, 2004, for a survey).

entrepreneur, profit is then

$$\pi^F = 2[\beta p - (w+s)\lambda - k] - c; \quad \beta > 1. \tag{2}$$

This equation reflects two costs and two benefits compared to informality. $c$ is a fixed cost of formality, which we refer to as a registration cost, though it may also be interpreted as a lump-sum tax. The second cost is that for each unit of labour employed, in addition to the market wage, a firm with formal status must pay the amount $s$, which, for most of our analysis we refer to as the cost of providing social benefits; but we can alternatively interpret $w+s$ as a government-imposed minimum wage for formal firms. Note that the $\lambda$-coefficient applies to an entrepreneur regardless of whether formal or informal status is adopted.

On the benefit side, it is assumed that, per package of labour and capital inputs, a formal firm earns $\beta p$ of revenue, rather than the $p$ earned by a formal firm. This may be because formal firms gain superior access to public goods such as public infrastructure, contract enforcement and property rights. It may also be the result of formal firms being able to sell their output to the government sector, fetching a higher price there than the market level $p$. The other benefit assumed for formal status is that it is associated with larger size. An entrepreneur may prefer to keep an informal firm small so as to avoid attracting the attention of the authorities. Given our other assumptions, an entrepreneur who adopts formality would always prefer larger to smaller size. To be specific, we assume that an entrepreneur who adopts formality uses twice as much capital and labour as when he or she adopts informality.[5]

[5]Our assumptions are consistent with evidence of McKenzie and Sakho (2007) for Bolivia. They find that the cost of registering as formal, in the form of then facing a liability for taxes,

We denote the opportunity cost for a $\lambda$-entrepreneur of running a firm in the industry we consider by $\bar{\pi}$, which, for simplicity, is assumed independent of $\lambda$. This is consistent with entrepreneurial ability being specific to the industry, in which case, as in Rauch (1991), $\bar{\pi}$ could be the wage an entrepreneur would make as a member of the labour force.[6]

Using (1) and (2), in Figure 1 we plot the loci for $\pi^I = \bar{\pi}$, $\pi^F = \bar{\pi}$, and $\pi^F = \pi^I$ in $(\lambda, p)$-space. The relative heights and slopes of these loci depend on whether $c \gtreqless \bar{c}$ and $s \gtreqless \bar{s}$, where

$$\bar{c} = 2(\beta - 1)k + (2\beta - 1)\bar{\pi}; \tag{3}$$

$$\bar{s} = (\beta - 1)w. \tag{4}$$

Thus, four cases are depicted in the figure (for simplicity, we disregard the possibility of equality, i.e., of $c = \bar{c}$ or $s = \bar{s}$) . In each case all three loci cut the $p$-axis at positive values of $p$, while in cases (i), (ii) and (iv) all three loci intersect in the positive quadrant at $\lambda = \lambda^*$ and $p = p^*$, where

$$\lambda^* = \frac{2(\beta - 1)k + (2\beta - 1)\bar{\pi} - c}{2[s - (\beta - 1)w]}; \quad p^* = \frac{2sk + (2s + w)\bar{\pi} - cw}{2[s - (\beta - 1)w]}. \tag{5}$$

[Figure 1]

For $(\lambda, p)$-combinations above (below) the $\pi^F = \pi^I$ locus, $\pi^F > (<)\ \pi^I$. Above (below) the $\pi^I = \bar{\pi}$ locus, $\pi^I > (<)\ \bar{\pi}$, and similarly for the $\pi^F = \bar{\pi}$ locus. Hence,

may be more than outweighed by the increase in customer base that is associated with the firm being able to issue tax receipts.

[6] Alternatively, we might have assumed that $\bar{\pi}$ is increasing in $\lambda$, i.e., that entrepreneurial ability in the industry we analyze has value in other industries. This would complicate the analysis (e.g., the loci in Figure 1 would then diminish in slope as $\lambda$ increases) without having a substantive effect on the results.

formality is chosen for $(\lambda, p)$-combinations above both the $\pi^F = \pi^I$ and $\pi^F = \bar{\pi}$ loci. Informality is chosen for $(\lambda, p)$-combinations below the $\pi^F = \pi^I$ locus but above the $\pi^I = \bar{\pi}$ locus. For $(\lambda, p)$-combinations below both the $\pi^F = \bar{\pi}$ and $\pi^I = \bar{\pi}$ loci an entrepreneur stays out of the industry. If $c >(<) \bar{c}$ we refer to $c$ as 'large' ('small'), and similarly for $s$.[7] In each panel of the figure the regions in which formality (informality) is chosen are denoted by F(I), and those in which the entrepreneur stays out are denoted by 0.

For informality to occur, at least one of the costs of formality, the registration cost $c$ and the social cost $s$, must be large. We shall not discuss any further case (iii), in which each of these costs of formality is small and so no entrepreneurs choose informality. For most of our analysis shall focus on case (i) in the figure, with $c$ small and $s$ large. Since case (ii), with both $c$ and $s$ large, is equivalent to case (i) for $\lambda > \lambda^*$, the results for (ii) are a subset of those for (i), and so we do not consider them separately. Case (iv), however, is associated with informality occurring under quite different conditions to cases (i) and (ii) and so will be discussed separately (see Section 5).

Evidence on the size of the registration cost $c$ is provided by Djankov et al. (2002). They find, across 85 countries at various levels of development, an average official registration cost of 47% of annual per capita GDP. From (3), in our model the critical value $\bar{c}$ is greater than the opportunity cost $\bar{\pi}$, possibly substantially so (depending on the values of $\beta$ and $k$). Although our model is static, $\bar{\pi}$ may be interpreted as a representation of the present value of the alternative earnings stream; but the registration cost is a one-off payment. These considerations sug-

[7] This simplifies the language, but it must be borne in mind that $\bar{c}$ and $\bar{s}$ are only large or small relative to the right-hand sides of (3) and (4), respectively.

gest that the assumption underlying case (i), that $c < \bar{c}$, may typically obtain in practice. However, Djankov et al. also find that the official registration cost is much higher in some countries, being above 100% of annual per capital GDP in 9 of their sample, the figure rising as high as 460% for the Dominican Republic. This suggests that a group of countries may exist with $c > \bar{c}$, that is, for which case (ii) or (iv) may apply.

From (4), $s > (<)\ \bar{s}$ as $(w+s)\lambda/\beta > (<) w\lambda$, that is, as, for a $\lambda$-entrepreneur, the unit labour cost for formality exceeds that for informality. Thus, in cases (i) and (ii), formality involves higher unit labour costs, while for cases (iii) and (iv) the reverse is true. If the productivity benefit $\beta$ is sufficiently large, case (iii) or (iv) will therefore obtain. Consider, however, the conditions that might make unit labour costs lower for formality than formality. These may relate to formal unit labour costs being relatively high or informal unit labour costs being relatively low.

Formal unit labour costs will be relatively high if formal firms pay rents to workers. Teal (1996) finds strong evidence of rent sharing with workers by firms in formal manufacturing in Ghana, rents being significantly associated with the presence of unions, with public and (partial or total) foreign ownership, and with higher profits. He concludes that the 30% differential between formal and informal sectors conjectured by Lewis (1954) is a substantial underestimate. It is also suggested by Collier and Gunning (1999) that in some African countries governments may generate rents for (formal) firms on the implicit understanding that the firms will pay high wage rates to workers. Thus, in our model satisfaction of the condition that $s > \bar{s}$ may be linked to $\beta$ exceeding unity. Informal unit labour costs may be relatively low if informal firms provide benefits that cost little but are greatly

valued by workers. For example, informal firms may provide greater flexibility of hours, and family members may be employed without an explicit wage being paid (see World Bank, 2007, on such behaviour in Latin America). Finally, note that in their survey of evidence on African manufacturing, Bigsten and Soderbom (2006) summarize evidence that wages are significantly greater in larger firms and that this can only partly be attributed to differences in worker characteristics.[8]

# 3 The Informal-Formal Mix

Let $\lambda_{\min}$ and $\lambda_{\max}$ denote the respective lowest and highest values of $\lambda$ in the pool of entrepreneurs. If only minimal entrepreneurial skills are required in the industry there may be many more (potential) entrepreneurs in the pool than would be needed to satisfy the demand for the product. Given heterogeneity of abilities, this implies that $\lambda_{\max}$ may be so large that we can treat it as infinite. A binding finite upper bound on $\lambda$ may exist, however, because the entrepreneurial skills that are needed in the industry are relatively scarce. We focus on this case because, as we show, a similar effect on the results will obtain even if the entrepreneurial skills are not scarce, but a complementary input is rationed. We discuss the example of a binding constraint on the aggregate supply of capital below.

In Figure 1 a horizontal line at any given price $p$ enables us to see the range of $\lambda$-values associated with informality, formality, or staying out of the industry. The

[8]For simplicity, we do not allow for the costs of capital being different for informal than for formal firms. Assuming that capital is more expensive for informal firms because of their lack of access to formal sources of finance, so that they use moneylenders, it becomes less likely that the appropriately amended version of (4) will be satisfied. However, informal firms may get the capital from family and friends. In Africa the interest rate for such loans is typically at or close to zero. For example, in Ghana, La Ferrara (2003) finds a real interest rate of approximately zero among kin groups.

configuration of behaviour that obtains depends on the distribution of $\lambda$ among the pool of entrepreneurs, and from this we can determine the output supply curve, which can be broken into informal and formal components. Combining this with the demand curve we can find what combination of informal and formal production obtains in equilibrium.

We adopt the following notation. $\pi^I = \pi^F$ for

$$\lambda(p) = \frac{(2\beta - 1)p - k - c}{w + 2s} \equiv \bar{\lambda}(p); \tag{6}$$

$\pi^I = \bar{\pi}$ for

$$\lambda(p) = \frac{p - k - \bar{\pi}}{w} \equiv \lambda^I(p), \tag{7}$$

and $\pi^F = \bar{\pi}$ for

$$\lambda(p) = \frac{\beta p - k - (c + \bar{\pi})/2}{w + s} \equiv \lambda^F(p). \tag{8}$$

We shall also refer the inverse functions of (6)-(8). The $p$ for which, for a given $\lambda$, $\pi^I = \pi^F$, is written $\bar{p}(\lambda)$; the $p$ for which, for a given $\lambda$, $\pi^I = \bar{\pi}$, is written $p^I(\lambda)$; the $p$ for which, for a given $\lambda$, $\pi^F = \bar{\pi}$, is written $p^F(\lambda)$.[9]

We express aggregate supply and demand as per entrepreneur in the pool. The cumulative density of $\lambda$ is denoted by $G(\lambda)$, and the supply of output by informal and by formal firms, per entrepreneur, by $q_s^I(p)$ and $q_s^F(p)$, respectively, where total supply per entrepreneur is $q_s(p) = q_s^I(p) + q_s^F(p)$. For brevity, we shall henceforth omit the phrase 'per entrepreneur.'

To keep the language simple the results will be stated on the assumption that the $\lambda$-distribution is continuous, but they can easily be written more generally.

[9] These values of $p$ are $\bar{p}(\lambda) = \frac{1}{2\beta - 1}\left[(w + 2s)\lambda + k + c\right]$; $p^I(\lambda) = w\lambda + k + \bar{\pi}$; and $p^F(\lambda) = \frac{1}{\beta}\left[(w + s)\lambda + k + \frac{1}{2}(c + \bar{\pi})\right]$.

For the following proposition, which characterizes the supply curves for formal, informal and total supply, we assume that $\lambda_{\min} < \lambda^* < \lambda_{\max}$, as illustrated in Figure 1(i).

**Proposition 1** *Consider case (i) with* $\lambda^* \in [\lambda_{\min}, \lambda_{\max}]$. *Then for* $p \in [0, p^F(\lambda_{\min}))$ *supply is zero. For* $p \in [p^F(\lambda_{\min}), p^*]$ *supply is only formal and is increasing in* $p$. *For* $p \in (p^*, \bar{p}(\lambda_{\max})]$ *supply is a mix of formality and informality and is increasing in* $p$; *formal supply is increasing in* $p$ *for all* $p \in (p^*, \bar{p}(\lambda_{\max})]$, *while informal supply increasing in* $p$ *for* $p \in (p^*, p^I(\lambda_{\max})]$, *but decreasing in* $p$ *for* $p \in (p^I(\lambda_{\max}), \bar{p}(\lambda_{\max})]$. *For* $p \in (\bar{p}(\lambda_{\max}), \infty)$ *supply is all formal and is constant.*

The proposition can be explained by considering progressively higher prices in the context of Figure 1(i). Raising $p$ from zero, formal entry will become profitable for an entrepreneur for whom $\lambda = \lambda_{\min}$ at $p$ just above $p^F(\lambda_{\min})$ (where $\pi_F = \bar{\pi}$ cuts $\lambda = \lambda_{\min}$). For $p \in [p^F(\lambda_{\min}), p^*]$, there will be formal entry by all entrepreneurs with $\lambda \leq \lambda^F(p)$ (where $p$ cuts the $\pi_F = \bar{\pi}$ locus). In this price range,

$$\begin{aligned} q_s^F(p) &= 2\beta G[\lambda^F(p)]; \\ q_s^I(p) &= 0. \end{aligned} \tag{9}$$

For $p \in (p^*, \bar{p}(\lambda_{\max})]$, formal entry occurs by entrepreneurs with $\lambda \leq \bar{\lambda}(p)$, that is, where $p$ cuts the $\pi^I = \pi^F$ locus. Informal entry occurs by entrepreneurs with $\bar{\lambda}(p) < \lambda \leq \min\{\lambda^I(p), \lambda_{\max}\}$, that is, between where $p$ cuts the $\pi^I = \pi^F$ locus and where it cuts the $\pi^I = \bar{\pi}$ locus, but only up to $\lambda = \lambda_{\max}$, after which, since the entire pool of entrepreneurs has entered, the only effect of further increases in price is to induce switches from informality to formality. Thus, for the price range

$p \in (p^*, \bar{p}(\lambda_{\max})]$,

$$\begin{aligned} q_s^F(p) &= 2\beta G[\bar{\lambda}(p)]; \\ q_s^I(p) &= G[\min\{\lambda^I(p), \lambda_{\max}\}] - G[\bar{\lambda}(p)]. \end{aligned} \tag{10}$$

Total supply rises until price reaches $\bar{p}(\lambda_{\max})$, when all entrepreneurs in the pool are producing formally.

Thus, above the minimum price that induces entry, low prices are associated with formal production only; intermediate prices are associated with a mix of formal and informal production; and high prices are associated with formal production only.

Intuitively, Proposition 1 can be explained in terms of the balance, for any $\lambda$ and $p$, of the productivity benefit of formality, the greater labour costs of formality - especially so for higher $\lambda$ - and the larger size (and so, potentially, ability to make profit) of a formal firm. If price exceeds $p^F(\lambda_{\min})$, but is no greater than $p^*$, entry is profitable for an entrepreneur with sufficiently low $\lambda$, that is, sufficiently low labour costs. The impact of a low $\lambda$ is greater for formality, which uses more labour, than for informality, and is reinforced by the so, also taking into account the productivity benefit $\beta$ of formality. Given that $c < \bar{c}$, the registration cost $c$ of formality is not so great as to prevent this outcome.

When price is raised above $p^*$, both formality and informality are profitable for a range of $\lambda$-values. But, within this range, for entrepreneurs with relatively high $\lambda$, the greater labour costs of formality make it less profitable than informality. Thus, the entrepreneurs with a relatively high $\lambda$ choose informality, while those with a relatively low $\lambda$ can profitably bear the higher labour requirements of formality.

As price is raised yet further, entrepreneurs with a higher $\lambda$ can enter profitably, but for those near the higher end of this range, informality is more profitable than formality because the saving in labour costs outweighs the benefits of greater size and of the productivity benefit $\beta$. However, as price rises further, and more entrepreneurs choose formality, eventually all the pool of entrepreneurs chooses formality.

To explore the role of informality further, assume that distribution of $\lambda$ is uniform, with mean $\Lambda$ and upper and lower bounds $\Lambda \pm \delta$ ($\Lambda > \delta > 0$), where $\Lambda - \delta < \lambda^* < \Lambda + \delta$. Thus, $\Lambda - \delta = \lambda_{\min}$ and $\Lambda - \delta = \lambda_{\max}$ and the cumulative density function is $G(\lambda) = (\lambda - \Lambda + \delta)/2\delta$ for $\lambda \in [\Lambda - \delta, \Lambda + \delta]$.

**Proposition 2** *Consider case (i) with a uniform distribution of* $\lambda$, $\lambda \in [\Lambda - \delta, \Lambda + \delta]$, *where* $\Lambda - \delta < \lambda^* < \lambda + \delta$. *For* $p \in (p^*, \hat{p})$ *the option of informality has a positive impact on output, but for* $p \in (\hat{p}, \bar{p}(\lambda_{\max}))$ *it has a negative impact on output, where* $p = \hat{p}$ *solves* $(2\beta - 1)G[\bar{\lambda}(p)] + 1 = 2\beta G[\lambda^F(p)]$.

This is illustrated in Figure 2, where $\bar{q}_s^F(p)$, which is shown by the thick line, denotes what supply would be if informality were somehow ruled out entirely. This supply curve is upward sloping until $p = p^F(\lambda_{\max})$, where all the pool of entrepreneurs are operating formally, and supply becomes vertical. Given, however, that the option of informality is open to entrepreneurs, this option is exercised as price $p$ is increased through $p^*$, the aggregate supply curve being $q_s^F(p) + q_s^I(p)$, as shown.

[Figure 2]

Starting from $p = p^*$, $q_s^F(p) + q_s^I(p)$ is an upward-sloping straight line until $p = p^I(\lambda_{\max})$. In this price range some entrepreneurs would choose formality if

informality were not available, but choose informality when it is available, and the formal supply curve $q_s^F(p)$ is to the left of $\bar{q}_s^F(p)$. But some entrepreneurs can make a profit with informality who could not make a profit from formality and so are drawn into production. The net result is that the total supply curve $q_s^F(p) + q_s^I(p)$ is to the right of $\bar{q}_s^F(p)$.

For $p$ above $p^I(\lambda_{\max})$, as $p$ increases there is a steady switch of entrepreneurs from informality to formality, but no additional entrepreneurs available to produce informally. Formal supply increases at the same rate as in the lower price range, but informal supply falls. The net result is that aggregate supply $q_s^F(p) + q_s^I(p)$ slopes up more steeply than in the lower price range. This occurs for prices up to $\bar{p}(\lambda_{\max})$, which corresponds to where $\pi^F + \pi^I$ cuts $\lambda = \lambda_{\max}$ in Figure 1(i). Above this price, the supply curve is vertical, supply being all formal.

As can be seen from Figure 2, aggregate supply with informality is less than it is without informality for the price range $p \in (\hat{p}, \bar{p}(\lambda_{\max}))$. To justify intuitively why aggregate supply may be smaller when informality is possible, consider what happens at price $p = p^F(\lambda_{\max})$. In Figure 1(i), this is the price at which $\pi^F = \bar{\pi}$ cuts $\lambda = \lambda_{\max}$. It can be seen from Figure 1(i) that at this price, if informality were excluded, the entire pool of entrepreneurs would produce formally. When informality is possible, however, the less efficient in the pool choose informality. Since all are producing in both these scenarios, but a $\lambda$-entrepreneur produces less with informality than with formality, it follows that output is smaller when informality is possible.

An implication is that as $p$ rises, $\phi$, the proportion of active firms that is formal, falls from unity and then rises back to unity. More precisely, we have the following.

**Corollary 1** *Under the conditions specified in Proposition 2,* $\phi = 1$ *for* $p \in [p^F(\Lambda - \delta), p^*]$*;* $0 < \phi < 1$ *with* $d\phi/dp < 0$ *for* $p \in (p^*, p^I(\Lambda + \delta)]$*;* $0 < \phi < 1$ *with* $d\phi/dp > 0$ *for* $p \in (p^I(\Lambda + \delta), \bar{p}(\Lambda + \delta))$*; and* $\phi = 1$ *for* $p \in [\bar{p}(\Lambda + \delta), \infty)$*.*

Combining this characterization of the supply side with a demand curve for output, we have that if demand is sufficiently tight (so that $p \leq p^*$) or sufficiently loose (so that $p \geq \bar{p}(\lambda_{\max})$) all supply will be formal; but for intermediate demand there will be a mix of formality and informality. However, if the assumption that $\lambda^* \in [\lambda_{\min}, \lambda_{\max})$ does not hold, the non-monotonicity result disappears. If $\lambda_{\min} \geq \lambda^*$ then at the lowest price for which positive output occurs, some supply will be informal; and if price is raised far enough, informality will disappear. Alternatively, if $\lambda_{\max} < \lambda^*$, informality is never chosen. The non-monotonicity result also disappears if $\lambda_{\max}$ is so large that, given that demand is not indefinitely large, market equilibrium is always to the left of the vertical portion of the supply curve.

A distinction can be drawn here. If the industry is such that a high level of entrepreneurial skills is needed, and these skills are in short supply relative to derived demand for them, then we shall have the case described in Proposition 1. If demand is large enough, market equilibrium will be on the vertical stretch of the supply curve, and there will be no informality, though if demand is not so large there may be some informality. Alternatively, if, only minimal entrepreneurial skills are required then $\lambda_{\max}$ may be large, in which case we shall see some informality even if demand is at a relatively high level.

If Proposition 2 is reconsidered in terms of employment, rather than output, we obtain the following.

**Corollary 2** *Under the conditions specified in Proposition 2, for* $p \in (p^*, \hat{p}^l)$ *the*

*option of informality has a positive impact on employment, but for* $p \in (\hat{p}^l, \bar{p}(\lambda_{\max}))$ *it has a negative impact on employment, where* $p = \hat{p}^l$ *solves* $[\lambda^F(p)]^2 = \{[\bar{\lambda}(p)]^2 + (\Lambda + \delta)^2\}$.

For this case Figure 2 can be amended so that employment, rather than output, is on the horizontal axis, but still with the goods price on the vertical axis. In place of each upward-sloping straight-line stretch in Figure 2, both formal derived labour demand and total derived labour demand are composed of curve segments that are upward-sloping because a higher price induces more entrepreneurs to produce both formally and *in toto*. But the curve segments have diminishing slopes because the entrepreneurs drawn into production are progressively less able, and so have progressively greater labour requirements. With these amendments, a result similar to Proposition 2 is obtained, though $\hat{p}^l$, the critical value of $p$, is higher than $\hat{p}$ (but below $p^F(\lambda_{\max})$). The availability of the option of informality may cause aggregate employment to fall because entrepreneurs who would otherwise choose formality, with its higher level of employment, can, for a range of parameter values, obtain a higher profit from informality.

As we have noted, if there is a binding constraint on the aggregate supply of an input that is complementary to entrepreneurship, the effect is similar to that of exhaustion of the supply of entrepreneurs. Suppose, for example, there is a limit on the aggregate amount of capital available (though no limit on the supply of entrepreneurs) and that any rationing of capital is efficient in the sense that an entrepreneur with a lower $\lambda$ gets priority over one with a higher $\lambda$. Assume that as $p$ is raised the capital constraint binds first at price $p_a$, where $p_a > p^*$, i.e., where, in the absence of the capital constraint there would be both formal and informal

production.

As $p$ rises above $p_a$, some entrepreneurs wish to switch from informality to formality and others wish to enter informally. Since in equilibrium those preferring formality have a lower $\lambda$ than those preferring informality, the constraint will bind for those preferring informality. The increased capital requirements of the entrepreneurs who wish to switch to formality will be satisfied, preventing additional entrepreneurs from entering informally and taking away capital from the least efficient of those who were already producing informally. The supply of formal output will rise by the same amount as it would in the absence of the capital constraint, but informal supply will fall. Aggregate supply will rise, but by less than it would in the absence of the capital constraint.

As $p$ rises further, this process will continue until $p = p_b$ is reached, at which all informal supply disappears, with all the supply of capital being used by firms with formal status. For any rises in $p$ above $p_b$ the capital constraint prevents more entrepreneurs from entering production. Thus, for $p \geq p_b$ there is no informal supply and the aggregate (formal) supply curve is vertical.

# 4 Comparative Statics

For a uniform distribution of $\lambda$, consider the effects of variation of parameter values for a given value of $p$. These are given in Table 1, which shows only the effects on output, but those on the corresponding employment terms are all of the same sign as those in the table. The parameters are shown in the columns and output in the rows. The column denoted $w^I$ refers to a change in $w$ with $w + s$ held constant. The first three rows show effects on output when $p \in (p^*, p^I(\Lambda + \delta)]$, and

the second three rows for $p \in (p^I(\Lambda+\delta), \bar{p}(\Lambda+\delta))$. The former range is referred to as 'low $p$' and the latter as 'high $p$.'

[Table 1]

In both price ranges formal output is increasing in the productivity benefit parameter $\beta$ and decreasing in the cost parameters $w$, $s$, $k$ and $c$. It is unaffected by small variations of opportunity cost $\bar{\pi}$ because entrepreneurs who choose formality earn profits in excess of $\bar{\pi}$. It increases if there is a rise in the informal wage cost $w^I$, which, at the margin of choice, causes a switch of entrepreneurs from informality.

For low $p$, informal output is decreasing in the cost parameters $w$, $w^I$ and $k$. Variation of parameters, $s$, $\beta$ and $c$, which impact directly on profits under formality, has effects on informal output of the opposite sign to those on formal output. Because the entrepreneur with the highest $\lambda$ of those in informal production just breaks even, a rise in the opportunity cost $\bar{\pi}$ has a negative effect on informal output.

The effects on informal output for high $p$ are perhaps more interesting. In this price range all entrepreneurs in the pool are producing (or all of the supply of complementary factors is exhausted), and, as a result, although formal output is unconstrained, informal output is a residual. The effects of parameter changes on informal output are therefore the opposite in sign to those for formal output. Thus, an increase in any of the formal cost parameters $w$, $s$, $k$ or $c$ has a positive effect on informal output - even though two of these parameters, $w$ and $k$, are also informal cost parameters. An increase in $w$ or $k$ has a direct negative effect on formal output with, at the margin of choice between formality and informality, some

entrepreneurs switching to informality, which, being smaller scale, is less affected by the cost increases. Because informality is a residual activity, and informal profit exceeds $\bar{\pi}$ for all firms producing, there is no counteracting shift from informality to inactivity. However, an increase in $w^I$, because it affects only informal production, causes a marginal switch from informality to formality, reducing informal output.

From (5) we find that both $\lambda^*$ and $p^*$ are increasing in $w$, $k$, $\beta$, and $\bar{\pi}$, and decreasing in $s$, $w^I$ and $c$. Using this result together with the equations for $\bar{p}(\lambda)$, $p^I(\lambda)$ and $p^F(\lambda)$, the effects of parameter changes on the loci in Figure 1 can be determined. Thus, the effect on the behaviour of each entrepreneur, and on the formal and informal supply curves, can be found, though, for brevity, we exclude this analysis here. We end our examination of comparative statics by discussing the impact of informality on the cyclical behaviour of the economy.

**Lemma 1** *Consider case (i) with a uniform distribution of* $\lambda$, $\lambda \in [\Lambda - \delta, \Lambda + \delta]$, *where* $\Lambda - \delta < \lambda^* < \lambda + \delta$. *Informality is a built-in stabilizer for* $p \in (p^I(\lambda_{\max}), p^F(\lambda_{\max}))$, *but a built-in destabilizer for* $p \in (p^*, p^I(\lambda_{\max})]$ *and* $p \in [p^F(\lambda_{\max}), \bar{p}(\lambda_{\max})]$.

In Figure 2, if the demand curve cuts both $q_s^F(p)$ and $q_s^F(p) + q_s^I(p)$ in the range $p \in (p^I(\lambda_{\max}), p^F(\lambda_{\max}))$, a small vertical variation in demand would cause a smaller change in output in the presence of informality than if it were prohibited. In this sense informality is a built-in stabilizer. Alternatively, however, suppose that the intersections with the demand curve occur outside this range, though above $p = p^*$ and below $p = \bar{p}(\lambda_{\max})$ the existence of informality adds to output instability. If the analysis is amended to relate to employment, rather than output, as discussed in Section 3, the hierarchy of slopes shown in Figure 2 for any given

$p$ is maintained and so the same conclusions apply with respect to employment as to output.[10]

# 5 Welfare

To examine the impact of informality on welfare, assume that the demand curve is downward-sloping and intersects the supply curve $\bar{q}^F(p)$ (i.e., when informality is ruled out) at $p = \bar{p}^F$, where $\bar{p}^F \in (p^*, \bar{p}(\lambda_{\max}))$. Thus, in Figure 2, the demand curve passes above point Z and below point Y. If the intersection were outside this range the introduction of the option of informality would have no effect: all production would be formal. We shall consider the components of welfare: consumer surplus, aggregate social benefits and aggregate profit,[11] and also note the impact on the amounts of 'good' and 'bad' jobs.

Consider first consumer surplus. In Figure 2, if the demand curve passes between points Y and X ($\bar{p}^F \in (\tilde{p}, \bar{p}(\lambda_{\max}))$), the option of informality causes the equilibrium total output to fall and price to rise. Hence, consumer surplus falls. If, however, the demand curve passes between points X and Z ($\bar{p}^F \in (p^*, \tilde{p})$), the option of informality causes equilibrium total output to rise and price to fall, so that consumer surplus rises.

Turning to total social benefits, the effect of the option of informality is clear-cut. Any downward-sloping demand curve passing between Z and Y cuts $q^F(p)$

[10] Some empirical evidence on this issue is provided by Loayza and Rigolini (2006). They find that in the short run, informal employment is counter-cyclical for the majority, but not all, developing countries.

[11] Since formal profit is calculated with the registration cost $c$ (a transfer) netted out, $c$ is also a component of welfare. But since any statements about the impact of the option of informality on aggregate social benefits also apply, qualitatively, to the impact on aggregate registration costs, we do not consider the latter separately.

to the left of where it cuts $\bar{q}^F(p)$. Therefore, the option of informality causes aggregate social benefits to fall; that is, it causes a reduction in the number of good jobs. Since there is informal production, however, bad jobs are created. This leads to the question: what is the effect on the total number of jobs?

Although Figure 2 has output on its horizontal axis, we have already noted how it may be transformed by having employment on this axis. Then, corresponding to the point X, which occurs at $p = \hat{p}$ in output-price space, there is a point at which the total amount employment is the same with and without the option of informality. This occurs at $p = \hat{p}^l$, where $\hat{p} < \hat{p}^l < p^F(\lambda_{\max})$. Thus, if $\bar{p}^F < \hat{p}^l$, the option of informality causes the total number of jobs to rise. The number of good jobs falls, but this is outweighed, numerically, by the number of bad jobs created. If, however, $\bar{p}^F > \hat{p}^l$, not only does the number of good jobs fall, but the number of bad jobs created is insufficient to offset this fall.

Because the equilibrium price is generally different with and without the option of informality, expressions comparing aggregate profit under each scenario are intractable. We note, however, that, depending on the specific demand curve, the option of informality can change aggregate profit in either direction.[12]

[12] As a limiting case, suppose first that demand is horizontal. Then for all $p > p^I(\lambda_{\max})$ all available entrepreneurs produce (or all complementary factors are used up). Since the option of formality is chosen by some, and this has no effect on the price received by the others, the availability of this option must have a positive effect on aggregate profit. Alternatively, suppose $p \in (p^*, p^I(\lambda_{\max})]$. Using (2), (5), (6) and (8), the option of informality for this price range changes aggregate formal profit by $\Delta\pi^F = -\int_{\bar{\lambda}(\bar{p}^F)}^{\lambda^F(\bar{p}^F)} 2[\beta p - \lambda(w+s) - k - c/2]d\lambda < 0$. Using (1) and (7), aggregate informal profit rises from zero to $\Delta\pi^I = \int_{\bar{\lambda}(\bar{p}^F)}^{\lambda^I(\bar{p}^F)} (p - \lambda w - k)d\lambda > 0$. Hence, $\Delta\pi^F + \Delta\pi^I = \frac{[(2s+w)\bar{\pi}+(p-p^*)z](p-p^*)[s-(\beta-1)w]}{(2s+w)(s+w)w}$. Since $p > p^*$ by assumption, $\Delta\pi^F + \Delta\pi^I > 0$ in this case.

If, however, we consider the alternative limiting case of a vertical demand curve, we can show by numerical examples that $\Delta\pi^F + \Delta\pi^I$ may be negative. If, for example, $k = c = \bar{\pi} = 0$, $w = 1$, $s = 0.5$, $\beta = 1.1$, $\Lambda - \delta = 0.1$ and $p = 1$, then the introduction of informality as an option reduces price to approximately 0.938 and $\Delta\pi^F + \Delta\pi^I \approx -0.015$.

These conclusions are summarized in our last proposition.

**Proposition 3** *Suppose that the demand curve is downward sloping and that, in the absence of informality, price is* $\bar{p}^F \in (p^*, \bar{p}(\lambda_{\max}))$. *Then the introduction of the option of informality (a) causes consumer surplus to rise if* $\bar{p}^F \in (p^*, \tilde{p})$, *but to fall if* $\bar{p}^F \in (\tilde{p}, \bar{p}(\lambda_{\max}))$; *(b) reduces aggregate social benefits; (c) reduces the number of good jobs and increases the number of bad jobs, with the total number of jobs rising if* $\bar{p}^F < \hat{p}^l$ *but falling if* $\bar{p}^F > \hat{p}^l$; *and (d) may affect aggregate profit in either direction.*

The impact of informality is to redistribute welfare; but it is also to redistribute profit. Consider the price ranges of Figure 2 again. If $\bar{p}^F \in (p^*, \tilde{p})$, then introduction of the option of informality causes entrepreneurs to produce (informally) who are less able than those producing (formally) in the absence of informality. Also, since there is a negative effect on price, the profits of the latter group of entrepreneurs fall. Hence, the effect of informality is to redistribute profit away from more able entrepreneurs towards less able ones. In a dynamic model this might be damaging for investment and growth prospects.

If $\bar{p}^F \in (\tilde{p}, p^F(\lambda_{\max}))$, introduction of the option of informality also brings less able entrepreneurs into production, but since there is a positive effect on price, there is also a positive effect on the profits of those producing formally in the absence of the informal option (the more able entrepreneurs). Finally, if $\bar{p}^F \in (p^F(\lambda_{\max}), \bar{p}(\lambda_{\max})]$, there is no effect on the number of entrepreneurs producing, but the positive effect on price again causes the more able entrepreneurs to gain higher profits. Thus, it is possible that the existence of the option of informality can improve investment and growth prospects.

# 6 Alternative Assumptions

In this section we consider the effects of three changes in assumptions, beginning with the definition of informality. It is possible in the model that some entrepreneurs at the high end of the $\lambda$-distribution will operate informally using more labour than those at the low end who operate formally, though the latter will nonetheless be using twice as much capital. This occurs if $\lambda_{\max} > 2\lambda_{\min}$. Suppose, however that we change our definition of informality to the condition that employment is less than a given size threshold $\lambda_0$. Assuming that $\lambda_0 < \lambda_{\max}$, our results must be amended.

In Figure 1(i) a vertical line would then be added at $\lambda = \lambda_0$. For $\lambda \in [\lambda_0, \lambda_{\max}]$, informality is not possible, but for $\{\lambda, p\}$-combinations on or above the $\pi^F = \bar{\pi}$ locus formality is chosen; that is, for points between the $\pi^F = \bar{\pi}$ and $\pi^F = \pi^I$ loci formality is now chosen, whereas in our analysis in previous sections informality was chosen. In Figure 2, the ZW segment of $q_s^F(p) + q_s^I(p)$ then does not extend as far as W. At $p = p^I(\lambda_0)$ a new segment slopes up parallel to WX: the restriction on informality caused by the employment condition has reduced the range of $\lambda$ for which informality occurs. However, this new segment does not extend as far as $\bar{q}_s^F(p)$. At $p = p^F(\lambda_0)$ another, steeper, segment begins, which meets $\bar{q}_s^F(p)$ at V: along this segment, as $p$ rises (a) there is a steady switch from informality to formality, as described previously, but also (b) entrepreneurs with $\lambda \in (\lambda_0, \lambda_{\max})$ switch from not producing to producing formally. These changes affect the details of our previous conclusions, but do not change their general character.

For Figure 1(iv), with $c > \bar{c}$ and $s < \bar{s}$, the latter inequality implies that, for a given entrepreneur, unit labour costs are less for formality than for informality.

In this case, as $p$ rises, the first entry, by the entrepreneurs with relatively low $\lambda$, is informal. This is because for such entrepreneurs labour usage is small, and so the unit labour cost (and size) advantage of formality is outweighed, in profit calculations, by the relatively large cost $c$ of registration (together with the cost of social benefit provision). As $p$ is raised further more entrepreneurs enter informally, but then a value of $p$ is reached above which the entrepreneurs with the lowest $\lambda$ prefer formality to informality. This is because formality is associated with greater production and so is favoured, relative to informality, by the higher $p$. When $p$ is raised far enough (through the intersection of all three loci) informality disappears, with all additional entry being formal. Indeed, if there were no entrepreneurs with $\lambda$ less than the level at which the loci intersect, there would be no informality at any price.

Consider now the implications of $s$ being provided for formal workers by the government with the government provision being funded by a profits tax on formal firms. If the tax were lump sum it could be regarded as already included in the registration cost $c$. Suppose, however, the tax is proportional, at rate $t$. Formal profits are then given by

$$\pi^F = (1-t)[2(\beta p - w\lambda - k) - c].$$

Assuming that $t < 1/2$, it is found that there is a critical value, $\tilde{c}$ of $c$:

$$\tilde{c} = 2[k(\beta - 1) + \pi\beta] - \frac{\pi}{1-t}.$$

If $c > \tilde{c}$ then the slopes and positions of the $\pi^I = \pi^F$, $\pi^F = \bar{\pi}$ and $\pi^I = \bar{\pi}$ loci are as

in case (iv) of Figure 1, with informality occurring for low $\{\lambda, p\}$-combinations, but not otherwise. If $c > \tilde{c}$ then the solution is as in case (iii) of Figure 1, informality not occurring for any $\{\lambda, p\}$-combinations.

Finally, we come to the representation of entrepreneurial heterogeneity. Our analysis is based on the assumption that differences in entrepreneurial ability are manifested in terms of differing labour costs. Heterogeneity might instead have been represented by a multiplicative coefficient on output that differs across entrepreneurs. Thus, we would have $\pi^I = \alpha p - w - k$ and $\pi^F = 2(\alpha\beta p - w - s - k) - c$, where $\alpha > 0$ and a higher $\alpha$ denotes a more able entrepreneur. A necessary condition for informality to be chosen is found to be that

$$s > (\beta - 1)(w + k) + [(2\beta - 1)\bar{\pi} - c]/2.$$

If this condition is satisfied, informality is chosen for $\alpha \in [(w + k + \bar{\pi})/p, (w + 2s + k + c)/(2\beta - 1)p)$ and formality is chosen for $\alpha \in [(w + 2s + k + c)/(2\beta - 1)p, \infty)$. As $p$ is raised informal supply falls, but formal supply increases such that total supply rises. The analysis is then similar to that for case (ii) in Figure 2, and therefore is also covered by our analysis of case (i).[13]

# 7 Conclusion

This paper has developed a framework for analyzing the mix of formality and informality in an industry with entrepreneurs of heterogeneous ability and has

[13] Corresponding to Figure 1, a figure can be drawn in $(\alpha, p)$-space. The curves (in the positive quadrant) for $\pi_I = \pi_F$, $\pi_I = \bar{\pi}$ and $\pi_F = \bar{\pi}$ are downward-sloping and non-intersecting. For $s > (\beta - 1)(w + k) + [(2\beta - 1)\bar{\pi} - c]/2$, $\pi_I = \pi_F$ is the highest and $\pi_I = \bar{\pi}$ the lowest.

considered some implications for welfare. It has focused on a particular application of this framework, with an aggregate constraint on the supply of entrepreneurs, or on a complementary factor, that binds for sufficiently high levels of product demand. At these levels of demand the impacts of informality being available as an option to entrepreneurs are that output and consumer surplus are reduced and, although some bad jobs are created, the number of good jobs is reduced by a larger amount. Nonetheless, the option of informality causes a redistribution of profits to more able entrepreneurs, which may be beneficial in the long run.

At lower levels of demand (though still enough to generate both formality and informality in equilibrium), the opposite effects obtain. The existence of the option of informality for entrepreneurs results in higher output and consumer surplus, and the number of bad jobs created outweighs the number of good jobs lost; but it also causes a redistribution of profits to less able entrepreneurs. These conclusions would also hold if there were no binding constraints on the supply of entrepreneurs or complementary factors, though for many industries in poorer countries the scenario with constraints may be the relevant one.

We also find that informality can be a built-in stabilizer or destabilizer, depending on the level of demand, but the ranges for which each of these conclusions apply do not match those for which informality raises or lowers supply. It is only in an intermediate range of demand that it is a built-in stabilizer. Outside this range, provided there is some informality in equilibrium, it acts as a built-in destabilizer.

# References


[1] Ahsan, Ahmad, and Carmen Pages (2007), Are all labor regulations equal? Assessing the effects of job security, labor dispute, and contract labor laws in India, World Bank Policy Research Working Paper 4259.

[2] Amaral, Pedro S, and Erwan Quintin (2006), A competitive model of the informal sector, Journal of Monetary Economics, 53, 1541-53.

[3] Banerji, Arup, and Sanjay Jain (2007), Quality dualism, Journal of Development Economics, 84, 234-250.

[4] Bennett, John, and Saul Estrin (2007), Informality as a stepping-stone: entrepreneurial entry in a developing economy, IZA Discussion Paper No 2950.

[5] Bigsten, Arne, and Mans Soderbom (2006), What have we learned from a decade of manufacturing surveys in Africa? World Bank Research Observer 21, 241-265.

[6] Bosch, Mariano (2007), Job creation and job destruction in the presence of informal labour markets, London School of Economics, mimeo.

[7] Bourguignon, Francois (2005), Development strategies for more and better jobs, paper presented at the conference 'Help Wanted: More and Better Jobs in a Globalized Economy,' Washington DC.

[8] Collier, Paul, and Jan Willem Gunning (1999), Explaining African growth performance, Journal of Economic Literature, 37, 12, March, 64-111.

[9] Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Sialnes, and Andrei Shleifer (2002), The regulation of entry, Quarterly Journal of Economics, 117, 1-35.

[10] Fields, Gary (2005), A guide to multi-sector labor market models, Cornell University, mimeo.

[11] Fortin, Bernard, Nicolas Marceau and Luc Savard (1997), Taxation, wage controls and the informal sector, Journal of Public Economics, 66, 293-312.

[12] Goldberg, Nina, and Pinelopi Pavcnik (2003), The response of the informal sector to trade liberalization, Journal of Development Economics, 72, 463-496.

[13] Ihrig, Jane, and Karine S. Moe (2004), Lurking in the shadows: the informal sector and government policy, Journal of Development Economics, 73, 541-557.

[14] Jovanovic, Boyan (1982), Selection and the evolution of industry, Econometrica, 50, 649-670.

[15] La Ferrara, Eliana (2003), Kin groups and reciprocity: a model of credit transactions in Ghana, American Economic Review, 93, 1730-1751.

[16] Lewis, W. Arthur (1954), Economic development with unlimited supplies of labour, Manchester School, 28, 139-191.

[17] Loayza, Norman V., and Jamele Rigolini (2006), Informality trends and cycles, World Bank Policy Research Working Paper 4078.

[18] Lucas, Robert E. (1978), On the size distribution of business firms, Bell Journal of Economics, 9, 508-523.

[19] Maloney, William (2004), Informality revisited, World Development, 32, 1159-1178.

[20] McKenzie, David, and Yaye Seynabou Sakho (2007), Does it pay firms to register for taxes? The impact of formality on firm profitability. IZA Discussion Paper No 3179.

[21] Parker, Simon C. (2004), The Economics of Self-Employment and Entrepreneurship, Cambridge University Press, Cambridge, UK.

[22] Paula, Aureo de and Jose Scheinkman (2006), The informal sector, UCLA Department of Economics Working Paper.

[23] Rauch, James E. (1991), Modelling the informal sector informally, Journal of Development Economics, 35, 33-47.

[24] Schneider, Friedrich (2006), Shadow economies of 145 countries all over the world: what do we really know? Working Paper, Johannes Kepler University of Linz.

[25] Schneider, Friedrich, and Dominik Enste (2000), Shadow economies: size, causes, and consequences, Journal of Economic Literature, 38, 77-114.

[26] Straub, Stephane (2005), Informal sector: the credit market channel, Journal of Development Economics, 78, 299-321.

[27] Teal, Francis (1996), The size and sources of economic rents in a developing country manufacturing labour market, Economic Journal, 106, 963-976.

[28] Walther, Richard and Ewa Filipiak (2007), Vocational Training in the Informal Sector, Research Department, Agence Française de Développement.

[29] World Bank (2007), Informality: Exit and Exclusion, World Bank, Washington D.C.

# Appendix: Proofs

## Proposition 2

For $p \in (p^*, p^F(\lambda_{\max})]$, $\bar{q}_s^F(p) = 2\beta G[\lambda^F(p)]$, so that, using (8), $\bar{q}_s^{F\prime}(p) = \beta(2\beta - 1)/\delta(w+2s)$. For $p \in (p^*, \bar{p}(\lambda_{\max})]$, $q_s^F(p) = 2\beta G[\bar{\lambda}(p)]$, so that, using (6), $q_s^{F\prime}(p) = \beta^2/\delta(w+s)$. Using (4), it follows that $\bar{q}_s^{F\prime}(p) > q_s^{F\prime}(p)$. Since output is on the horizontal axis in Figure 2, $\bar{q}_s^F(p)$ has the lesser slope.

For $p \in (p^*, p^I(\lambda_{\max})]$, $q_s^I(p) = G[\lambda^I(p)] - G[\bar{\lambda}(p)]$. Using (6) and (7), we obtain $q_s^F(p) + q_s^I(p) - \bar{q}_s^F(p) = \left[2\beta(\frac{(2\beta-1)p-k-c}{w+2s} - \frac{\beta p-k-(c+\pi)/2}{w+s}) + \frac{p-k-\pi}{w}\right]/2\delta \equiv x/2\delta$. Then $dx/dp = \{2s[s-(\beta-1)w] + sw + w^2 + 2w^2\beta(\beta-1)\}/(2s+w)(s+w)w$, which, given (4), is positive. However, substituting $p = p^*$ into $q_s^F(p) + q_s^I(p) - \bar{q}_s^F(p) = x/2\delta$ and using (3) and (4), we find that $q_s^F(p^*) + q_s^I(p^*) - \bar{q}_s^F(p^*) > 0$. Hence, $q_s^F(p) + q_s^I(p) - \bar{q}_s^F(p) > 0$ for $p \geq p^*$.

For $p \in (p^I(\lambda_{\max}), \bar{p}(\lambda_{\max})]$, $q_s^I(p) = 1 - G[\bar{\lambda}(p)]$. Using (6), $d[q_s^F(p) + q_s^I(p)]dp = (2\beta-1)^2/2\delta(w+s)$. Using (4), $d[q_s^F(p) + q_s^I(p)]dp < \bar{q}_s^{F\prime}(p)$. The proposition follows.

## Corollary 2

For $p \in (p^*, p^I(\lambda_{\max})]$, we have from Proposition 2 that $q_s^F(p) + q_s^I(p) > \bar{q}_s^F(p)$. Since labour productivity is higher with formality than informality, it follows that employment with informality, $l_s^F(p) + l_s^I(p)$, is greater than employment without, $\bar{l}_s^F(p)$. (Similar reasoning applies to all $p$ up to $\hat{p}$.) For $p \in (p^I(\lambda_{\max})), \bar{p}(\lambda_{\max})]$, $l_s^F(p) + l_s^I(p) = \frac{1}{2\Lambda}\int_{\Lambda-\delta}^{\bar{\lambda}(p)} 2\lambda d\lambda + \frac{1}{2\Lambda}\int_{\bar{\lambda}(p)}^{\Lambda+\delta} 2\lambda d\lambda$. Using (6), this is increasing in $p$. However, for the upper part of this $p$-range, $p \in [p^F(\lambda_{\max})), \bar{p}(\lambda_{\max})]$, $\bar{l}_s^F(p)$ is constant. Since $l_s^F(p) + l_s^I(p) = \bar{l}_s^F(p)$ at $p = \bar{p}(\lambda_{\max})$, we have that for $p \in [p^F(\lambda_{\max})), \bar{p}(\lambda_{\max}))$, $l_s^F(p) + l_s^I(p) < \bar{l}_s^F(p)$. The proposition follows, with the value of $\hat{p}^l$ defined by the equation $l_s^F(p) + l_s^I(p) = \bar{l}_s^F(p)$.

Table 1 Comparative Statics

| | | $\beta$ | $k$ | $c$ | $\bar{\pi}$ | $w$ | $s$ | $w^I$ |
|---|---|---|---|---|---|---|---|---|
| low $p$ | $q_s^F$ | $+$ | $-$ | $-$ | $0$ | $-$ | $-$ | $+$ |
| | $q_s^I$ | $-$ | $-$ | $+$ | $-$ | $-$ | $+$ | $-$ |
| | $q_s^F + q_s^I$ | $+$ | $-$ | $-$ | $-$ | $-$ | $-$ | $-$ |
| high $p$ | $q_s^F$ | $+$ | $-$ | $-$ | $0$ | $-$ | $-$ | $+$ |
| | $q_s^I$ | $-$ | $+$ | $+$ | $0$ | $+$ | $+$ | $+$ |
| | $q_s^F + q_s^I$ | $+$ | $-$ | $-$ | $0$ | $-$ | $-$ | $-$ |

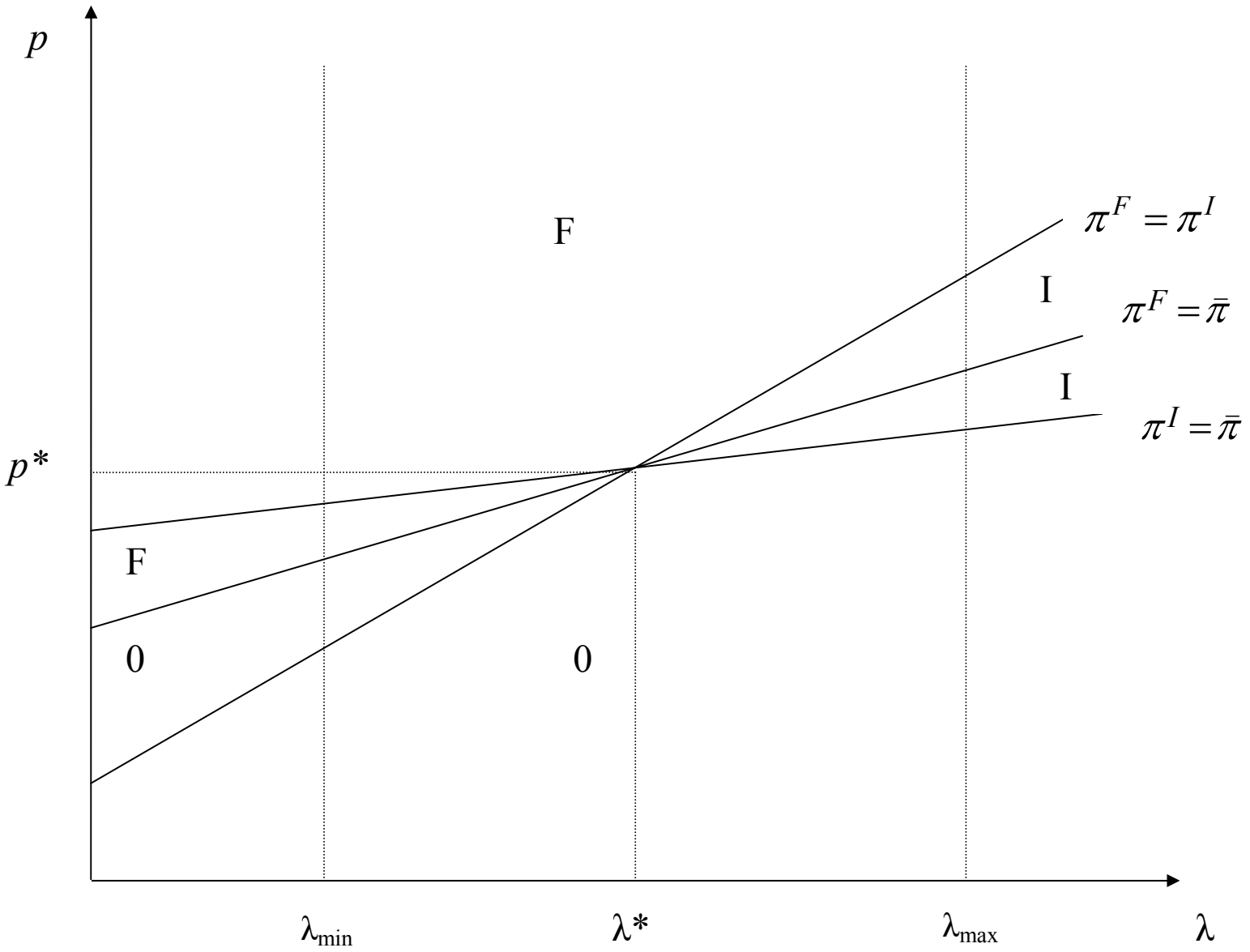


Figure 1(i) *c* small; *s* large

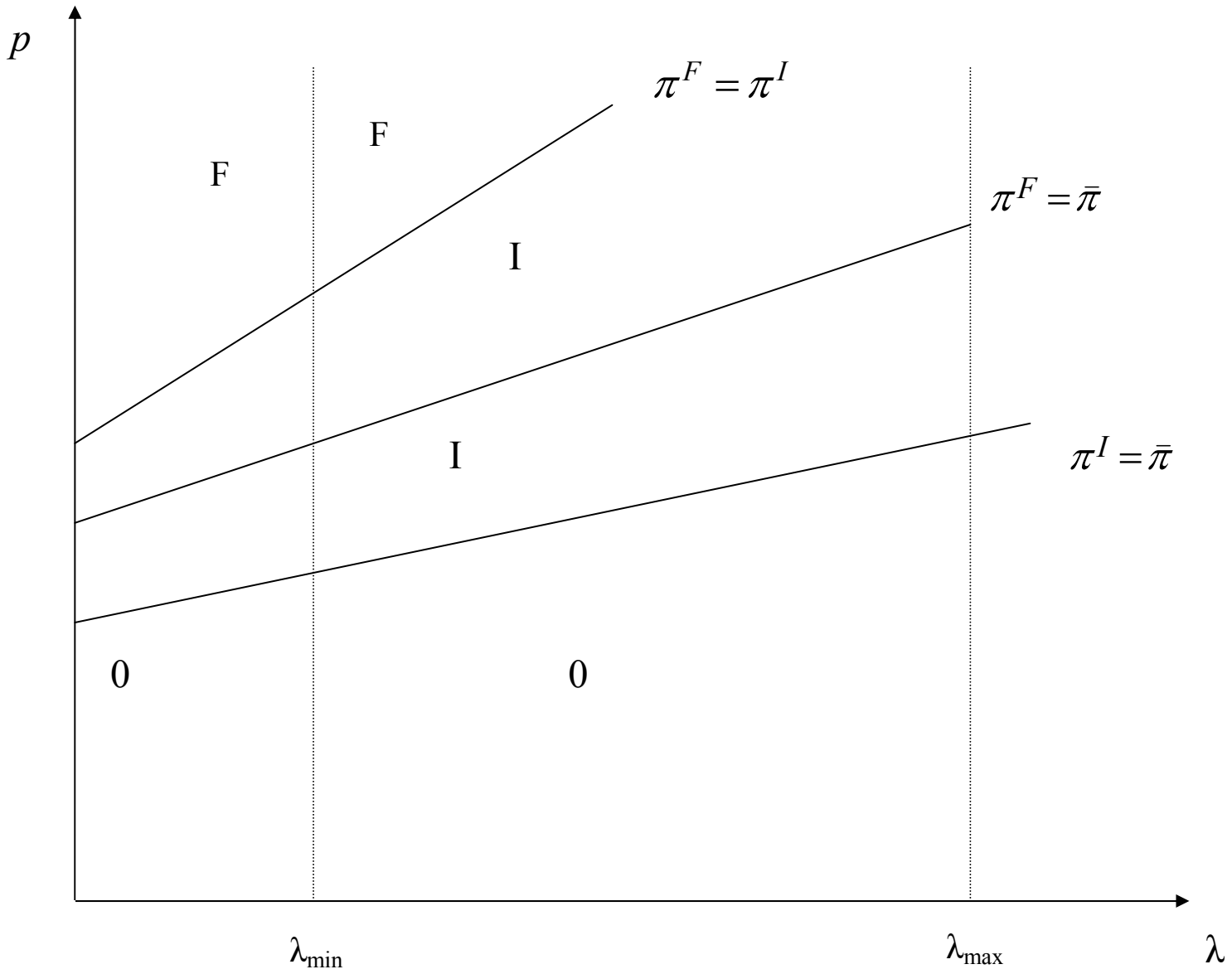


Figure 1(ii) *c* large; *s* large

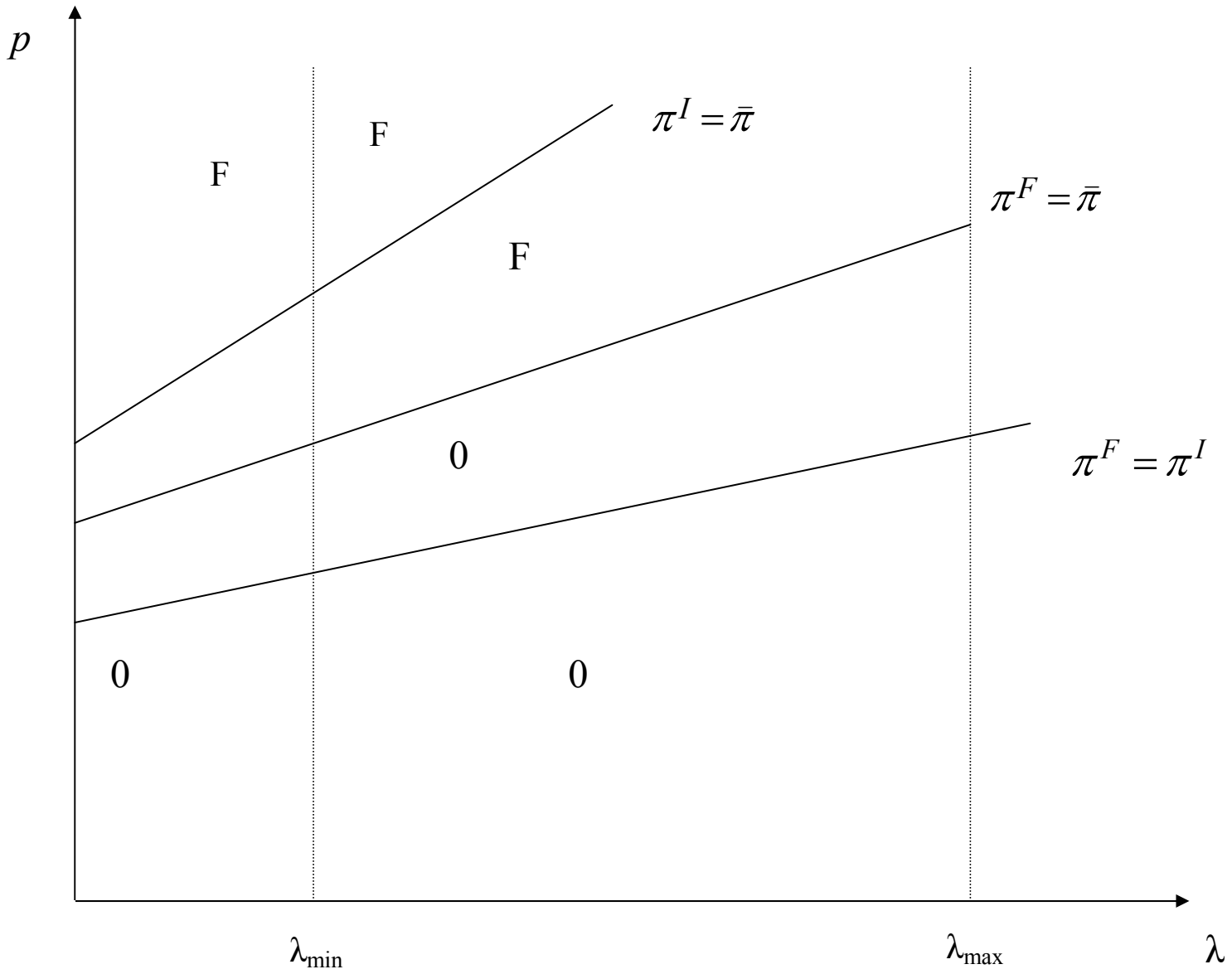


Figure 1(iii) *c* small; *s* small

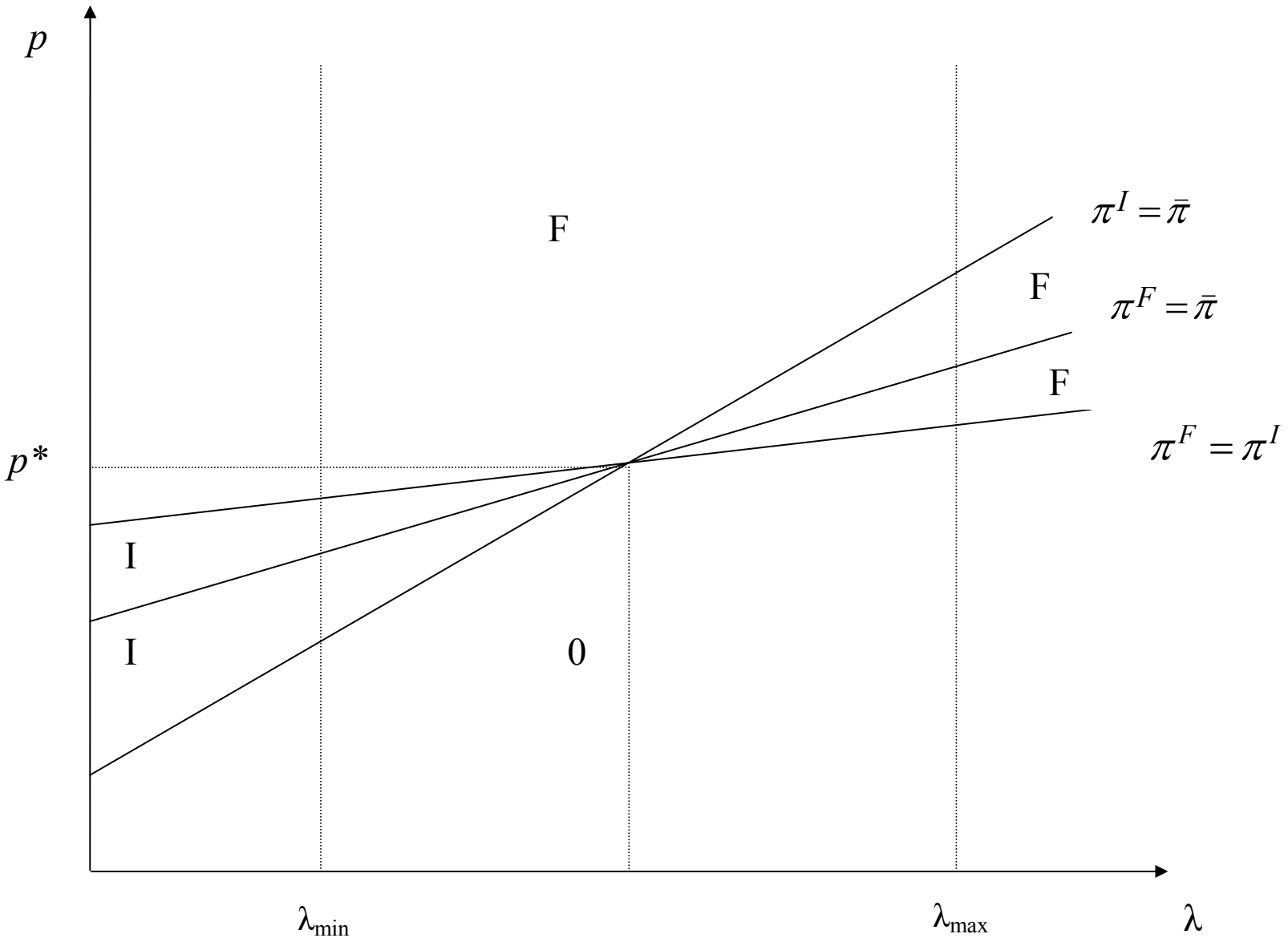


Figure 1(iv) *c* large; *s* small

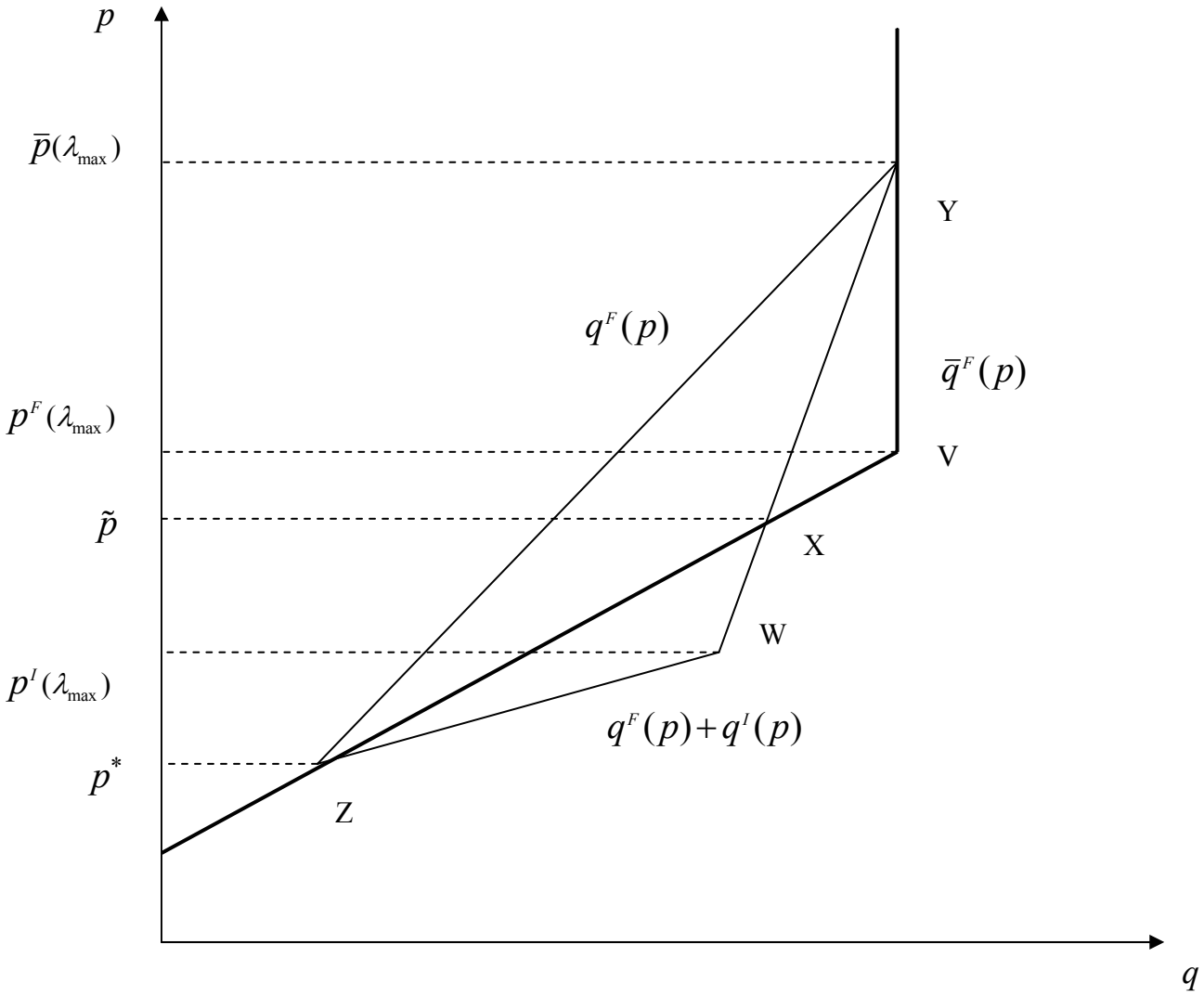


Figure 2 The Impact of Informality on Supply