Società italiana di
economia pubblica

XXIV
CONFERENZA

FINANCING SOCIAL POLICY IN THE PRESENCE OF INFORMALITY

EHTISHAM AHMAD, MICHAEL BEST

Financing Social Policy in the Presence of Informality*

Ehtisham Ahmad (ZEF, Bonn & Asia Research Centre, LSE)
Michael Best (STICERD and Department of Economics, LSE)†

31st March 2012

Abstract

We present a framework for the analysis of tax and benefit policy in countries with significant informality. Our framework allows us to jointly analyse the effects of various taxes and benefits on incentives for firms and workers to be informal and evade taxation. We find that payroll taxes and targeted minimum income guarantees targeted to households without formal employment are particularly harmful to labour formality and participation in the modern sector labour force. Conversely, Bismarkian benefits targeted to formal sector workers and basic benefits targeted to low income households represent the least distortionary way to redistribute. Attempts to use holes in the VAT to "protect" the poor are generally ineffective and open up avenues for rent seeking. We also find that a uniform value added tax and a corporate income tax represent the least distortionary way to raise revenues. The information generated from a simple VAT can be used, given an appropriately designed tax administration, to enhance the probability of detection of informal activities. Distributional issues are best handled by social policy and the personal income tax. Indeed, given the gainers and losers from tax reforms, social policies and intergovernmental transfers will be needed to ensure the political acceptance of the reforms. The precise mix of taxation and social policy will vary given different country characteristics and institutional structures.

1 Introduction

This paper presents a general framework for the analysis of public policy making in situations of informality. This is particularly important in developing countries as diverse as Mexico and Pakistan, where both the social policy design, as well as financing, represent "good intentions, but bad outcomes," using the insightful terminology of Santiago Levy (2008). However, the insights carry over also to EU countries in crisis, such as Greece, where informality and unemployment pose significant barriers to the sustainability of Bismarkian social protection mechanisms and constrain the effective operation of the main standard tax tools–including the corporate and income taxes, payroll contributions and the value added tax (VAT).

Our framework allows us to examine how various forms of taxation affect incentives for informality of workers and firms and how these feed into the overall labour productivity, efficiency and output of the economy. We build on existing work on Mexico by Levy and coauthors (Levy, 2008; Antón *et al.*, 2011), finding, as they do, that payroll taxes give rise to particularly unfavourable incentives for formality. What is more, we find that payroll taxes also depress incentives for participation in the modern sector labour force and that they reduce labour productivity and output. We also find that corporate income taxes reduce incentives for employing workers informally and hence increase productivity, as formal workers, but not informal workers, are expensable in determining a firm's corporate income tax (CIT) base. We also find that a uniform VAT preserves production efficiency, though it provides incentives to under-report sales. Incentives to under report sales depend on both the corporate income tax rate and the VAT rate and on their interaction, suggesting that raising revenue through a combination of the two instruments may be more effective than relying excessively on either one.

Indeed, the main advantage of a properly designed VAT is to generate information that could be used to tighten administration, and with positive results for the VAT and also income tax revenues. A VAT with exemptions and holes for the poor generally does little for the poor, becomes a source of rent seeking and loses revenues. Concern for the poor would then appropriately depend on social policy design, and the personal income tax would address income distributional issues. The balance across taxes would depend on country circumstances, and revenue needs–it is unlikely that raising the VAT indefinitely (e.g., to 27% as recently in Greece) would be a sensible approach in the presence of informality. The changing effects of tax reforms on resources at the disposal of governments at different levels, would require a joint assessment and modifications in intergovernmental transfers to ensure that states/provinces do not lose from the reforms. The issue of assignments of the major taxes to subnational governments could include simple "piggy-backed options" rather than the complex splitting of the base that is seen in the Indian sub-continent–these issues are not addressed in this paper.

On the social policy side we find that benefits that only formal workers are eligible for provide strong incentives for workers to be formal and also have positive effects on participation in the modern sector labour force, labour productivity and output. This makes some formal sector benefits an attractive instrument for redistribution purposes since formal workers' wages

are known and thus it will be feasible to effectively target lower income "formal" households. Conversely, targeting households without formal employment could have the effect of increasing informality if the level of the benefit is too high. If workers are able to simultaneously receive the benefit and take informal employment, but would lose the benefit if they become formal workers, then such a benefit would increase the attractiveness of being in the informal sector. Furthermore, such benefits could also discourage participation in the modern sector labour force and depress productivity and output.

This suggests that targeting benefits according to the employment status (formal or informal) of workers should be done cautiously. Notwithstanding this, benefits targeted to households with low income levels (independently of their employment status) can provide an attractive vehicle for redistribution, provided due care is taken in their design to take into account the incentives they present to low income households. Many countries could expand existing programmes for the poor (whether formal or informal) to compensate for adverse relative price changes, e.g., coming about from eliminations of zero-ratings and exemptions in the VAT.

The public finance literature on informality in developing countries has largely focused on how they should balance consumption taxes (usually the VAT) and trade taxes. Appropriately, the main issues here have been the ability of such taxes to reach the informal sector and the ability of governments to administer and enforce taxes, such as the VAT (see, for example Keen & Mintz, 2004; Emran & Stiglitz, 2005; Keen, 2008 and Boadway & Sato, 2009). This literature usually takes an optimal commodity taxation approach, building on the classical Ramsey framework and Atkinson & Stiglitz (1976) (an important exception that is relevant to our work is Dharmapala *et al.*, 2011). However, we think it is important to include income taxation and labour informality in the design of a desirable mix of instruments,given significant revenue requirements in the developing country context. Indeed we show that there are important interactions between the CIT and the VAT in the presence of informality that existing work has not highlighted.

The literature on informality relating to firms has tended to focus on registration decisions of the firms, and particularly on small scale microentrepreneurs (see, for example, La Porta & Shleifer, 2008; de Mel *et al.*, 2010 and McKenzie & Sakho, 2010). There is also a literature in developing countries focusing on how connected or not the formal and informal labour markets are (see, for example Bosch & Maloney, 2010 and Günther & Launov, 2012). In the literature, the closest work to our approach is work that combines firms and workers in modeling the incentives for informality such as Levy (2008) and de Paula & Scheinkman (2010), which we build on.

In sections (2) and (3) we present our framework, focusing first on workers and then on firms. Section (4) uses the framework to analyse the effect of changing the rates at which taxes are levied, the levels of benefits, the effectiveness with which benefits are targeted and the effectiveness of detection of tax evasion. Finally, section (5) discusses extensions for future work and section (6) concludes.

2   Workers

Societies are populated by households with diverse earnings potentials. To capture this feature in the simplest possible way, we will assume that there are two types of workers, with high and low ability, and that high ability workers are relatively scarce. Specifically, there is a mass 1 of high ability workers, and a mass $N > 1$ of low ability workers. Individuals also differ in their employment status. Some individuals operate entirely outside the modern sector. This may be because they are unemployed, or because they work in agricultural occupations, or are self employed in a manner that they are not included in the tax net. Either way, these workers are not registered to pay payroll (or income) taxes; and if they work, they are in occupations that are also outside the tax net. Furthermore, as they are unregistered workers, they do not receive benefits available only to formal workers.

Individuals who work in the modern sector may or may not be registered with the authorities. We refer to unregistered workers as informal workers. It is important to note that, following most of the literature on informality (see, for example, Levy, 2008; Kanbur, 2009) we do not equate informality with illegality. Informal workers may be illegal workers, but they may also be workers working on commission or be self employed, and as such may not be illegal. Formal workers receive a salary from an employer, are fully registered with the tax and social security authorities and are, hence, likely compliant with labour and tax laws.

Empirically, higher skilled and older individuals are less likely to be informal workers (see, for example, Perry *et al.*, 2007; Bosch & Maloney, 2010) and so for simplicity we will assume that all high skilled individuals are formal workers. The low ability individuals may be outside the modern sector, informal workers, or formal workers. This assumption is made only to keep the model tractable. It can be relaxed without losing any of the qualitative results here, but would add significantly to the complexity of the equations.

Note that we assume that workers are either completely informal or completely formal. That is, informal workers are not registered either for tax or for benefit purposes. In most countries this is not problematic as it would be difficult to lie to one authority and not to others, however, in many countries where the administrations of taxes and benefits are substantially separated, it may be possible, for example, for a worker not to pay payroll (or income) taxes, but to receive benefits that only registered workers are eligible for.

All high ability workers are equally productive, and they all receive a wage w_H. We will also assume that the high skilled labour market is competitive, so that both firms and workers take the wage as given. Low ability workers' wages will depend on their legal status as well as their skill level (as is discussed in more detail in section 3 below), with informal workers receiving w_I and formal workers receiving w_F.

All households are eligible to receive a universal benefit at a level b_U. This benefit is provided to all households, regardless of skill level/income or formal status. For example, Bolivia's universal pension (*Renta Dignidad*) can be thought of in this way. Second, formal sector workers (all high skilled workers and low skilled workers in the formal sector) are eligible to receive a

formal sector benefit b_F. An example of this kind of benefit is the provision of healthcare and insurance through Social Security (IMSS) in Mexico.

Third, there is a targeted transfer intended for low income households, b_L. However, the targeting of the program (through, for example, means testing) is imperfect and the government's screening technology only accurately identifies the low-skilled households with probability γ. This means that low ability households expect to receive γb_L while high income households will expect to receive $(1 - \gamma) b_L$. Imperfections in targeting may be technological in the sense that they stem from the difficulty of observing the eligibility status of a household, or they may stem from rent-seeking by the implementing officials who extract bribes from ineligible households in order to receive the benefit. In addition, if implementation is by lower levels of government than that responsible for the financing of the benefit (usually the center), there could be a possibility of generating agency problems.

Finally, there is a targeted minimum income guarantee (TMIG) targeted at low ability households who are not working in the formal sector (note that workers may simultaneously receive this benefit and work informally, though not all will) set at a level b_O. In many countries, a natural example of such a benefit is unemployment insurance, which is tied to an individual's formal employment sector. Indeed, here there is a wealth of evidence that individuals sometimes simultaneously receive benefits and work informally, and also that the level of such benefits affects the intensity with which individuals seek employment. In developing countries, an example of such a benefit is India's National Rural Employment Guarantee Act (NREGA), a workfare scheme giving publicly funded work guarantees to workers in rural areas, at least to the extent that it is possible for households to receive the NREGA work and simultaneously work informally as, say agricultural labourers. This would no longer be possible if the workers were to take formal work.

While high ability workers are all formal, we have not yet discussed the allocation of low ability workers to the formal and informal sectors. Our setup is not one that is equipped to explicitly deal with involuntary unemployment, so the allocation of low skilled individuals to their three possible occupations happens through the interplay of the willingness of the low ability households to work in the three sectors, the willingness of firms to hire formal and informal workers, and the incentives for these choices presented by the tax and benefit system. That said, the distinction between self-employment in the informal sector and unemployment is not particularly sharp and may not be policy relevant here (Rauch, 1991). While some informality is doubtless the result of the rationing of jobs in the formal sector, much recent work has shown that most informality is voluntary in the sense that it responds to incentives and therefore that one should focus on these incentives when analysing the allocation of households to these three sectors (Maloney, 2004; La Porta & Shleifer, 2008; Bosch & Maloney, 2010).

Low ability households vary in two respects. First, low ability households vary in the opportunity cost of working in a firm, either formally or informally. This opportunity cost may be the foregone earnings from subsistence agriculture or small-scale self employed activities, or it may be a fixed cost of going to work in the modern sector (for example, in cases where we think of b_O

	Ability Level	
	Low Ability	High Ability
Formal	$w_F + b_U + b_F + \gamma b_L + \mu_i$	$w_H + b_U + b_F + (1-\gamma)b_L$
Informal	$w_I + b_U + b_O + \gamma b_L$	N/A
Outside	$\varepsilon_i + b_U + b_O$	N/A

Table 1: Consumption Levels of Three Sorts of Households

as an unemployment benefit), or finally it could be the disutility from taking formal or informal work in the private sector, or something like a caste based barrier to gaining employment in the (in)formal sector as, for example might be the case for government employment schemes like NREGA in India. We denote the value of this outside option as ε_i which is assumed to be distributed among the population according to the well-behaved cumulative distribution function $K(\varepsilon)$.

Second, low ability households are assumed to have an idiosyncratic attachment to work in the formal sector, conditional upon working in the modern sector that influences their willingness to work informally. This idiosyncratic component may stem from variations in how households value the benefits available to them in the formal sector (see, for example Levy, 2008, chapter 3), or the peace of mind from knowing that the household is behaving in accordance with the law etc. This attachment to the formal sector is denoted by μ_i, which is distributed among the population according to the well-behaved cumulative distribution function $J(\mu)$ independently of ε_i. Individuals' expected utility levels will, therefore, be as summarised in table 1.

In order to establish how low ability workers choose between the three options open to them, we think of their occupational choice as a two stage problem. First, households choose whether to seek (in)formal employment or to stay out of the private sector labour force. Then, if the worker chooses to seek (in)formal employment, she chooses whether to seek formal or informal work. For simplicity, we will assume that households do not know their attachment to the formal sector μ_i when choosing whether or not to seek (in)formal employment, but rather have beliefs regarding their μ_i distributed (correctly) according to $J(\mu)$. Imagine, for example, that working in the modern sector requires traveling to the city to seek work. In this case, households may first decide whether to travel to the city and then once in the city, decide what sort of work to seek.

Solving backward, let us start in the second stage with the decision of whether to work formally or informally. An individual with a given level of μ_i receives utility of $U_{FL} = C_{FL} + \mu_i$, where $C_{FL} = w_F + b_U + b_F + \gamma b_L$ if they work as a formal worker, while their utility from being informal is simply their consumption $U_{IL} = C_{IL} = w_I + b_U + b_O + \gamma b_L$. Workers will choose to work in the formal sector if $U_{FL} > U_{IL}$. That is, they will choose to seek formal work if their attachment to the formal sector exceeds a critical value $\tilde{\mu}$.

$$\mu_i > w_I - w_F + b_O - b_F = \tilde{\mu}$$

This critical value of the attachment to the formal sector is determined by the relative wages that workers in the two sectors receive, but also by the difference between the benefit that they can receive as an informal worker, b_O, but would lose if they become a formal worker and the benefit that they will become eligible to receive as a formal worker, b_F. As we will see below in section 4, these benefits will have this direct effect on informality, but the benefits will also affect the equilibrium wage levels w_I and w_F, so that they also have an indirect effect on informality through wages. Overall, of the workers who seek (in)formal work in the modern sector, a fraction $J(\tilde{\mu})$ of them seek informal work, while the remaining $1 - J(\tilde{\mu})$ seek formal work.

Working back to the decision to enter the modern sector, workers anticipate that if they seek (in)formal work in the modern sector instead of remaining outside, their expected utility is

$$\mathsf{E}[U] = J(\tilde{\mu}) U_{IL} + [1 - J(\tilde{\mu})] (C_{FL} + \mathsf{E}[\mu|\mu > \tilde{\mu}]),$$

while if they stay out of the modern sector labour force, they will continue to receive the benefits b_O and b_U and also receive their reservation value from remaining outside the modern sector ε_i. Comparing these, individuals will choose to enter the private labour force if

$$\varepsilon_i < \mathsf{E}[U] - b_O - b_U = \tilde{\varepsilon}.$$

As with the decision between formality and informality, there is a critical value of the reservation value $\tilde{\varepsilon}$ and individuals with a small enough reservation value will join the modern sector while the remainder will stay out of the modern sector. The determinants of this critical value $\tilde{\varepsilon}$ are more complicated than those that determine $\tilde{\mu}$ as they include all the determinants of $\mathsf{E}[U]$ (the wages, w_I and w_F, the benefits b_F, b_L and b_O, the effectiveness of targeting γ and the distribution of μ, $J(\mu)$), but nevertheless, we will show in section 4 how the various policy instruments affect participation in the modern sector.

Summarising the total number of individuals in each group, there will be $N_O = N[1 - K(\tilde{\varepsilon})]$ households outside the modern sector, $N_I = NK(\tilde{\varepsilon})J(\tilde{\mu})$ informal workers, and $N_F = NK(\tilde{\varepsilon})[1 - J(\tilde{\mu})]$ formal workers. In models of endogenous occupational choice and informality the highest ability individuals become entrepreneurs in the modern sector and intermediate ability individuals become informal entrepreneurs and they do so in response to the returns to these occupations (see, for example de Paula & Scheinkman (2010) which builds on Rauch (1991)). We assume that workers cannot become entrepreneurs in the modern sector and vice versa so that the total number of workers is fixed and the total number of firms in existence is also fixed. However, we do allow for the interpretation of being outside the modern sector as being in informal self-employed activities, so we capture this margin to some extent. This amounts to assuming that the incentives to become an entrepreneur in the modern sector are not affected by the tax and benefit system. This may be the case, for example, if starting a firm requires an operating license which is only available to those with political connections, or if entry into a sector requires a large sunk cost and there are credit constraints in the economy.

3 Firms

There is a rich literature studying the compliance of firms in developing countries with regulations. For instance, de Mel *et al.* (2010) and McKenzie & Sakho (2010) study the decision of firms to register with the authorities finding, respectively, that demand for formality among microenterprises is low and that mid-sized firms with high ability managers profit from being formal while small firms do not. Studying informal employment, La Porta & Shleifer (2008) and Levy (2008) find that many firms simultaneously employ formal and informal workers and that the propensity to employ formal workers increases with the size of the firm.

Our paper seeks to provide a framework for thinking about the public finance questions of how to finance social policy when informality and evasion are pervasive. In the spirit of the recent literature using sufficient statistics for welfare analysis (see Chetty, 2009, for a review), what matters for our purposes is the overall size and composition of the labour force, and the overall degree of tax compliance of firms. We adopt a very simple treatment of the firm which allows us to derive sufficient statistics for welfare analysis in a straightforward manner. While a richer treatment of firm responses would add greater colour to our framework, it would also add a large amount of complexity without, we believe, qualitatively altering the sufficient statistics we derive.

Firms produce output using labour of both skill types. However, formal and informal low-skilled labour are not equally productive (though they are perfect substitutes) with formal, low-skilled labour being more productive. The firm's production function is

$$Y\left(L_E, H\right) = L_E^{\alpha_L} H^{\alpha_H},$$

where Y is output, H is the quantity of high skilled labour used, and the input $L_E = \theta L_F + (1 - \theta) L_I$ is effective low-skilled labour, an aggregate of formal (L_F) and informal (L_I) low-skilled labour. This formulation assumes that formal and informal low skilled workers are perfect substitutes but that $1 > \theta > 1/2$ so that formal, low ability workers are more productive than informal, low ability workers. We also assume that $\alpha_H > \alpha_L$ so that high ability workers are more productive than low ability workers.

Firms face three taxes. First, there is a payroll tax of τ_p on the wage bill for formal workers (of both skill levels). Second, there is a value added tax on the firms' sales (since in our simple framework the only input is labour and there are no intermediate goods, the VAT effectively approximates a final point sales tax). Finally there is a corporate income tax on profits (here we assume that the formal wage bill and payroll tax are expensable).

For the purposes of paying taxes, firms must report their sales and the details of their workforce to the tax authorities. However, enforcement is imperfect and so firms have the opportunity to misreport their sales and/or their workforce. If they do so, they face some probability that this under-reporting might be detected, in which case they face a fine for the tax they have evaded. In the case of their sales, if a firm reports sales of $p\hat{Y}$ and has actual sales

of pY, the probability of detection is denoted by λ_Y.

Specifically, we assume that the probability of detection is proportional to the fraction of sales that is unreported, $\lambda_Y = \min\left\{\eta\frac{Y-\hat{Y}}{Y}, 1\right\}$. The parameter η controls the strength of enforcement of taxes, with larger values of η being associated with better enforcement. In practice, η will be larger when the tax system generates, collects and uses information effectively. In particular, the VAT is the instrument most clearly linked to the magnitude of sales. Generally speaking, taxes (especially the VAT) with numerous exemptions and complicated rate structures will tend to lend themselves to weak enforcement both because they generate less useful information and because they provide numerous opportunities for evasion and corruption. If a firm's under-reporting is detected, the firm faces a fine of F_Y per dollar of unreported sales (the fine here can also represent the lost profits from legal proceedings, loss of reputation and clients and, even, imprisonment). Combining these elements, the firm's expected fine is given by $\mathsf{E}\left[F_Y\right] = \lambda_Y F_Y p\left(Y - \hat{Y}\right)$.

Similar, to under-reporting of sales, hiring informal workers amounts to under-reporting the low-skilled workforce and is also illegal as the payroll contributions would have been evaded. Firms face another fine if discovered employing informal workers. If a firm has L_F formal unskilled workers and L_I informal unskilled workers (for a total of $L = L_F + L_I$ low skilled workers), the expected fine is $\mathsf{E}\left[F_L\right] = \lambda_L F_L L_I$ where F_L is a fine per informal worker and λ_L is the probability of detection. Here, we assume that the probability of detection is simply $\lambda_L = L_I/L$, the proportion of low ability workers that are informal.

A firm's expected profits are therefore

$$\Pi = (1-\tau_c)\left\{(1-\tau_v)\,p\hat{Y} - (1+\tau_p)\left[w_H H + w_F L_F\right]\right\} + p\left(Y - \hat{Y}\right) - w_I L_I - \mathsf{E}\left[F_L\right] - \mathsf{E}\left[F_Y\right] \quad (1)$$

where τ_c is the corporate income tax rate, τ_v is the VAT rate (or sales tax, in the one-good example) and τ_p is the payroll tax rate. Firms choose the number workers of the three types that they hire, L_F, L_I and H, and the sales that they report $\hat{Y}$ in order to maximise expected profits (1). Sections 3.1 and 3.2 below analyse the firm's decision to (under)report sales and hire informal labour, which are central to our framework, while the less novel considerations of the total amount of labour to demand and output to produce, as well as the demonstration that an equilibrium exists in this framework are relegated to appendix (A).

3.1 Informal sales

The profit maximisation condition for reported sales $p\hat{Y}$ is that

$$p\left\{2\eta F_Y\left(1 - \frac{\hat{Y}}{Y}\right) - \left[1 - (1-\tau_c)(1-\tau_v)\right]\right\} \le 0 \quad (2)$$

and rearranging equation (2) we see that the firm's optimal ratio of reported to total sales is

$$\frac{\hat{Y}}{Y} = 1 - \frac{1}{2\eta F_Y}\left[1 - (1-\tau_c)(1-\tau_v)\right] \quad (3)$$

Sales under-reporting is higher the higher is the VAT rate and also higher the higher is the corporate income tax rate. This is natural as the base of both of these taxes is larger the larger are reported sales. So as the tax rate rises, so does the incentive to under-report sales. However, there is also an interesting interaction between the two taxes. The higher is one tax, the more distortionary are changes in the other tax. How responsive reported sales are to the corporate income tax rate $\left(\partial \left(\hat{Y}/Y\right)\partial \tau_c\right)$ depends on the VAT rate, and vice versa. That is:

$$\frac{\partial^2 \frac{\hat{Y}}{Y}}{\partial \tau_c \partial \tau_v} = \frac{1}{2\eta F_Y} > 0.$$

This means that these taxes should be considered in unison and that high levels of either tax will not only lead to high levels of evasion of that tax, but also increase evasion of the other tax. This, in turn suggests that high levels of either tax are undesirable and that a combination of lower, and similar levels of both taxes is preferable.

Equation (3) also says that sales under-reporting is decreasing in η, the tax enforcement parameter, meaning that the probability of detection of sales under-reporting affects revenues from both VAT and corporate income taxes. The enforcement parameter reflects not only the strength of the tax administration's capacity in the sense of the training of enforcement officers and optimal auditing strategies etc., but also the design of tax policy. Complex rate structures in either the VAT or the CIT with extensive use of exemptions and zero ratings provide opportunities for evasion and corruption that would be reflected in our framework as a lower η and hence a lower ratio of reported to total sales.

3.2 Informal Workers

Turning next to the labour market, we first note that the analysis of the labour market of low skilled workers can be divided into two parts. First, the analysis of the effective overall demand and supply of unskilled workers L_E and second, the composition of the unskilled workforce in terms of formal and informal workers. We are most interested in the prevalence of informality amongst the low skilled workers, so we present the analysis of the second part here and relegate to the appendix the demonstration of the clearance of the overall market for low-skilled labour. In particular, we define $f^D = L_F^D/L$ as the relative demand for formal, low-skilled workers, and $\beta = w_F/w_I$ as the relative wage of formal workers. Note that since formal and informal unskilled labourers are perfect substitutes, in the provision of effective low skilled labour, firms will only employ both formal and informal workers if the cost per effective unit of unskilled labour (which we will denote by $\tilde{w}_L$) is the same for formal as for informal workers. That is, if the firm employs both formal and informal unskilled workers, it must be the case that[1]

$$\frac{1}{\theta}\left[(1-\tau_c)(1+\tau_p)\beta w_I - F_L\left(1-f^D\right)^2\right] = \frac{1}{1-\theta}\left[w_I + F_L\left(1-f^D\right)\left(1+f^D\right)\right] = \tilde{w}_L. \quad (4)$$

---

[1] This can be seen by combining the first order conditions for profit maximising choices of L_I and L_F (equations (7) and (8) in the appendix).

We are most interested in the prevalence of informality amongst the low skilled labour force so we define $f^D = L_F^D/L$ as the relative demand for formal, low-skilled workers. See the appendix for the demonstration that there exists a $\tilde{w}_L$ that clears the market for effective low-skilled labour.

We will analyse the relationship between the formal wage premium β and the rate of formality f for a given informal wage w_I, since for any combination of w_F and f, w_I is determined by equation (4) and the level of $\tilde{w}_L$ that clears the market for effective low-skilled labour. Applying the implicit function theorem to (4)

$$\frac{\partial f^D}{\partial \beta} = -\frac{\frac{1}{\theta}(1-\tau_c)(1+\tau_p)}{2F\left[\frac{1-f}{\theta} + \frac{f}{1-\theta}\right]} w_I \quad < \quad 0.$$

So, as is intuitive, the demand for formal relative to informal workers is decreasing in the wage that must be paid to formal workers relative to the informal wage. Just as for the relative demand for formal labour we can use the equations in section 2 to write the relative supply of formal labour as

$$f^S = \frac{L_F^S}{NK(\tilde{\varepsilon})} = 1 - J(\tilde{\mu}) \tag{5}$$

which is increasing in the formal wage since

$$\frac{\partial f^S}{\partial \beta} = j(\tilde{\mu}) w_I > 0.$$

This means that if we plot the demand and supply curves in $\beta \times f$ space, the demand and supply curves will cross and the unique equilibrium wage satisfies both (5) and (4). This is illustrated in figure (1). Having set up the framework, we can now use it to study the impact that the various taxes and benefits have on the size and welfare of the various groups in society. Once we have specified the role of the government in our model, we can also study how this framework should inform policy choices by governments, a task we turn to next.

3.3   The Government

For simplicity we ignore the role of government in providing public goods and focus exclusively on its redistributive and revenue raising role, as has been the case in the classical optimal income taxation literature (see, for example Mirrlees, 1971; Saez, 2001). The government will receive revenues from its three tax bases, corporate income, sales, and wages. Total revenues will therefore be

$$R = \tau_c \left\{ p\hat{Y} - (1+\tau_p)[w_H + w_F N_F] \right\} + \tau_v p\hat{Y} + \tau_p [w_H + w_F N_F].$$

On the expenditure side, the government must finance the benefits it provides: the universal benefit b_U, the targeted transfer b_L, the benefit for formal workers b_F, and the targeted minimum
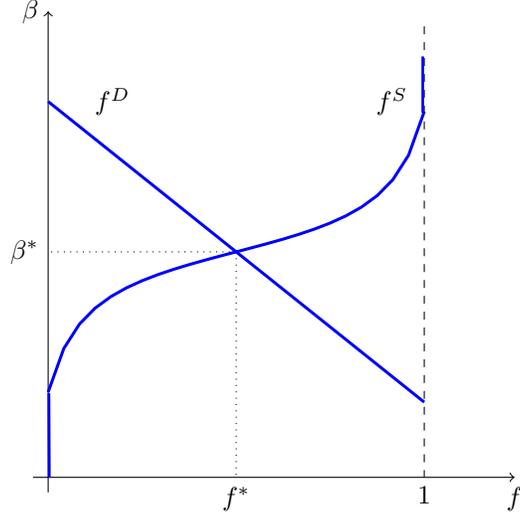
Figure 1: Labour Market for Low-skilled Labour

income guarantee b_O. Total expenditures are therefore

$$E = (1 + N) b_U + [1 - \gamma + \gamma N] b_L + b_F \{1 + N_F\} + b_O N_O.$$

We will ignore government debt and assume that the budget must be balanced, so that the government faces a budget constraint given by

$$B = E - R \leq 0.$$

Turning to the normative question of how these taxes and benefits affect welfare, we will assume that societal welfare W is simply the weighted sum of the welfare of the 5 groups in society (high skilled workers, 3 classes of low skilled workers and firms).[2] Therefore, the overall welfare in society is given by

$$W = g_{IL} N_I U_{IL} + g_{FL} N_F U_{FL} + g_{OL} N_O U_{OL} + g_H U_H + g_\Pi \Pi,$$

where the gs are welfare weights on the five groups in society (the 4 classes of households and firms) and Π is the profit of firms. We can also write this in Lagrangean form as

$$\tilde{W} = W - \mu B,$$

---

[2]We have deliberately not specified what these weights are. They may be exogenously given weights, or they may depend on the welfare of each group as in a traditional Atkinson welfare function, which has the property that the social welfare weights are decreasing in the individual's income at a rate determined by the society's inequality aversion.

where μ is the marginal cost of public funds (the Lagrange multiplier on the government's budget constraint) so that we can take account of how changes in the taxes and benefits affect welfare and the government budget. The government has at its disposal 3 taxes (τ_c, τ_p and τ_v), 4 benefits (b_U, b_L, b_O and b_F) and there are 2 implementation parameters (γ and η). So, we can differentiate the welfare function to see the effects of changing any of these 9 parameters. For any of these parameters $\rho \in \{\tau_c, \tau_p, \tau_v, b_U, b_L, b_O, b_F, \gamma, \eta\}$ the change in welfare resulting from a slight change in the parameter is

$$
\begin{aligned}
\frac{\partial \tilde{W}}{\partial \rho} &= g_{IL}\left[\frac{\partial N_I}{\partial \rho}U_{IL} + N_I \frac{\partial U_{IL}}{\partial \rho}\right] + g_{FL}\left[\frac{\partial N_F}{\partial \rho}U_{FL} + N_F \frac{\partial U_{FL}}{\partial \rho}\right] \\
&+ g_{OL}\left[\frac{\partial N_O}{\partial \rho}U_{OL} + N_O \frac{\partial U_{OL}}{\partial \rho}\right] + g_H \frac{\partial U_H}{\partial \rho} + g_\Pi \frac{\partial \Pi}{\partial \rho} - \mu \frac{\partial B}{\partial \rho}.
\end{aligned}
\tag{6}
$$

What matters is how each parameter affects the welfare of each group, how it affects the distribution of low-skilled workers among their three possible occupations as formal workers, informal workers, or in outside occupations, and how it affects the government's budget balance. Equation (6) specifies what empirical quantities we would want to know in order to estimate the effects of policy changes, much as in Ahmad & Stern (1991).

4   Effects of Taxes, Benefits and Implementation

In this section we analyse the impact of changes in the rates of taxes, the levels of benefits and the implementation parameters γ and η on the prevalence of informal employment, participation in the modern sector labour force and the wage premium of formal workers. These are interesting not only in themselves but also in that they would be the key ingredients in an empirical assessment of the desirability of any such reforms as they are the behavioural elasticities that represent sufficient statistics for welfare analysis in the sense of implementing equation (6) for the various taxes, benefits and administration efforts.

4.1   Taxes

As has been powerfully argued by Santiago Levy and coauthors (Levy, 2008; Antón *et al.*, 2011), payroll taxes can create a severe distortion by taxing the formal sector and effectively subsidising informality through the benefits available to informal workers that they do not contribute to. In other cases, minimum wage legislation may also act like a "tax" on formal sector employment, especially in cases where there is a fragmented labour market and cheaper foreign labour is available (either formally, through work visas, or through illegal migration). In our framework, this is also the case and, in fact, the effects are even worse as there is an additional negative effect on participation. As can be seen in figure 2, an increase in the payroll tax shifts the relative demand curve down and makes it flatter without directly affecting the relative supply
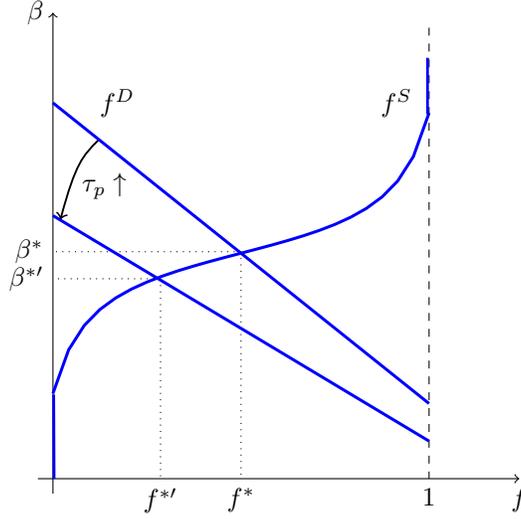
Figure 2: Increase in payroll tax

curve. [3] It does this because an increase in the payroll tax makes formal workers more expensive relative to informal workers (holding wages constant). The overall effect of this shift is to reduce formality and the formal sector wage premium as some of this extra cost of employing formal workers is passed on to the workers in the form of lower wages.

Turning to participation, the effect depends on what happens to $\tilde{\varepsilon}$. The decrease in formality (i.e. an increase in $\tilde{\mu}$) increases $\mathsf{E}[U]$,[4] but the decrease in the formal wage premium reduces $\mathsf{E}[U]$. However, the second effect dominates and the increase in the payroll tax reduces $\mathsf{E}[U]$, which means that $\tilde{\varepsilon}$ goes down and hence that participation in the modern sector goes down.[5] This decrease in participation will create a shortage of effective low-skilled labour and so the cost of effective, low-skilled labour $\tilde{w}_L$ will go up. This will shift the relative demand curve up and offset these effects somewhat, but these general equilibrium effects are second order and so cannot be strong enough to reverse the effects on formality, the formal sector wage premium,

---

[3]The demand curve shifts down because from equation 4 $\frac{\partial f^D}{\partial \tau_p} = -\frac{(1-\tau_c)\beta w_I}{2F_L\theta\left[\frac{1-f}{\theta}+\frac{f}{1-\theta}\right]} < 0$ and it becomes

flatter because $\frac{\partial^2 f^D}{\partial\beta\partial\tau_p} = -\frac{\frac{1}{\theta}(1-\tau_c)}{2F\left[\frac{1-f}{\theta}+\frac{f}{1-\theta}\right]}w_I < 0$.

[4]$\mathsf{E}[U]$ can be rewritten as $\mathsf{E}[U] = w_F + b_U + b_F + \gamma b_L + J(\tilde{\mu})\tilde{\mu} + [1 - J(\tilde{\mu})]\mathsf{E}[\mu|\mu > \tilde{\mu}]$ which means that $\frac{\partial\mathsf{E}[U]}{\partial\tilde{\mu}} = J(\tilde{\mu}) > 0$

[5]To see this, approximate the relative supply and demand curves around the equilibrium f and β by straight lines. The relative demand curve will have a slope of $-\frac{2F\left[\frac{1-f}{\theta}+\frac{f}{1-\theta}\right]}{\frac{1}{\theta}(1-\tau_c)(1+\tau_p)w_I}$ which we will denote ϕ and the relative supply curve will have a slope of $1/j(\tilde{\mu})w_I$ which we will denote ζ. Under this linear approximation, if the shift down in the demand curve is of size Δ, then the change in f, df is given by $-(\zeta - \phi)\Delta$ and the change in β, $d\beta$ is given by $-\zeta(\zeta - \phi)\Delta$. Since $f = 1 - J(\tilde{\mu})$, this means that the change in $\tilde{\mu}$ is $d\tilde{\mu} = -df/j(\tilde{\mu}) = (\zeta - \phi)\Delta/j(\tilde{\mu})$. Similarly, the change in the formal wage (holding the informal wage constant) is $dw_F = w_I d\beta = -\zeta w_I(\zeta - \phi)\Delta$. Finally, the total change in $\mathsf{E}[U]$ is $d\mathsf{E}[U] = dw_F + J(\tilde{\mu})d\tilde{\mu} \approx \left(-\zeta w_I + \frac{J(\tilde{\mu})}{j(\tilde{\mu})}\right)(\zeta - \phi)\Delta = \frac{J(\tilde{\mu})-1}{j(\tilde{\mu})}(\zeta - \phi)\Delta < 0$ demonstrating that the increase in the payroll tax reduces $\mathsf{E}[U]$.

and participation, only dampen them. Overall, the reduction in formality and the reduction in participation will lead to a fall in productivity and output in the economy. This may seem to form a compelling argument for the abolition of payroll taxes. However, this need not inevitably be the case.

Workers may perceive there to be a direct link between the payroll taxes they pay and the benefits they receive. Thus, they may perceive these taxes as contributions towards their eventual benefits (as in, for example, an individual retirement account). Under these circumstances, payroll taxes may actually generate an incentive for workers to want to be formal because of the insurance they will receive.

The analysis of the corporate income tax is the converse of that for the payroll tax. Because wages paid to formal workers can be deducted from a firm's corporate income tax base, a corporate income tax makes formal workers more attractive relative to informal workers. So, an increase in the corporate income tax rate will shift the relative demand curve up and make it steeper as in figure 3.[6] This will increase formality and raise the formal sector wage premium. The increase in formality is a decrease in $\tilde{\mu}$ which reduces $\mathsf{E}[U]$ while the increase in the formal sector wage premium β increases $\mathsf{E}[U]$, but as was the case with the payroll tax, the wage effect dominates and so the overall effect is to increase the expected utility of working in the modern sector $\mathsf{E}[U]$. This increases $\tilde{\varepsilon}$ and so increases participation in the modern sector labour force. Again, as was the case for the payroll tax, there will be offsetting second-order general equilibrium effects, but they will not be strong enough to reverse the direction of these effects.

This highlights the effective role that a well designed corporate income tax (CIT) can play in the tax system. Of course, we saw briefly in section 3.1, and as will be discussed below in section 4.3 in greater detail, the corporate income tax also generates incentives for firms to under-report their sales, and does so in a way that interacts with other taxes (particularly the VAT). Thus, one cannot just rely upon increasing the CIT exclusively to raise revenue.

Finally, in a simple, 1-sector framework, the VAT τ_v does not affect labour informality. In a richer model with multiple sectors, this may no longer be the case, depending on how complementary labour inputs are with intermediate inputs (this extension remains for further work as discussed in section 5). Also, one of the most appealing features of the VAT is that it preserves production efficiency (in the sense of Diamond & Mirrlees, 1971). In our framework with evasion the VAT still preserves production efficiency, only the reporting of sales and so this relative neutrality property of the VAT makes it an attractive instrument for raising revenues on a broad base with minimal distortions. Thus, extreme care should be taken in implementing departures from uniformity, as not only do these destroy production efficiency[7], these also create opportunities for evasion and rent seeking. However, as for the corporate income tax, a properly

---

[6]The demand curve shifts up because from equation 4 $\frac{\partial f^D}{\partial \tau_c} = \frac{(1+\tau_p)\beta w_I}{2F_L\theta\left[\frac{1-f}{\theta}+\frac{f}{1-\theta}\right]} > 0$ and it becomes steeper

because $\frac{\partial^2 f^D}{\partial \beta \partial \tau_c} = \frac{\frac{1}{\theta}(1+\tau_p)}{2F\left[\frac{1-f}{\theta}+\frac{f}{1-\theta}\right]} w_I > 0.$

[7]Indeed, Maurice Lalé, the father of the world's first VAT in France in 1954 describes exemptions as "the cancer of the VAT system" European Commission (2010, p.28)
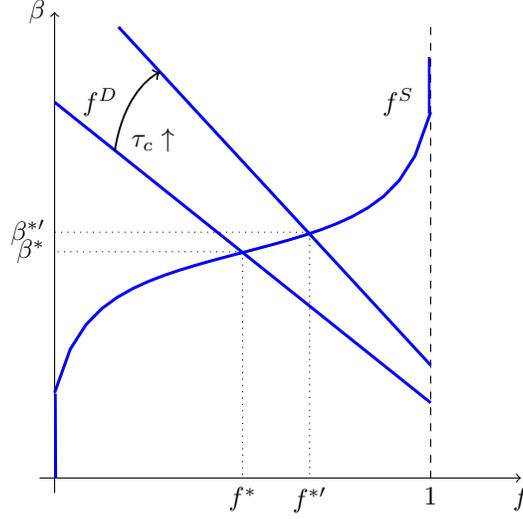
Figure 3: Increase in corporate income tax

designed VAT with minimal exemptions has the effect to reduce under-reporting of sales by firms, and should also raise CIT revenues without raising CIT rates.

4.2 Benefits

There are four benefits in our framework, the formal sector benefit b_F, the TMIG b_O, the targeted benefit b_L, and the universal benefit b_U. In this section we explore the comparative static effects of increasing these by a small amount, starting with the formal benefit b_F. Looking at equation (5), we can see that an increase in b_F will shift the supply curve down,[8] causing the formal sector wage premium β to fall and labour formality to rise as in figure 4. Since the relative demand is not perfectly inelastic, the fall in the wage premium will not be sufficient to offset the increase in consumption coming from the increase in b_F and so formal workers will be better off, making it more attractive to be a formal worker and thus increasing labour formality.

Work in the modern sector has become more attractive overall, raising $\mathsf{E}[U]$ and $\tilde{\varepsilon}$ and hence increasing participation in the modern sector, labour productivity and output. This increase in participation increases the supply of effective labour and so has the general equilibrium effect of reducing the cost of effective low-skilled labour $\tilde{w}_L$ which will shift the relative demand curve down, but this second order effect will not reverse the overall increase in labour formality. Since the authorities have reliable information on formal workers' wages, this analysis suggests that formal sector benefits are an appealing instrument to use to channel benefits to low-skilled households working in the formal sector. More accurate information availability will enhance possibilities of targeted or categorical benefits and that these will also act to encourage labour

---

[8]It shifts the supply curve down because $\frac{\partial f^S}{\partial b_F} = \frac{\partial f^S}{\partial \tilde{\mu}} \frac{\partial \tilde{\mu}}{\partial b_F} = j(\tilde{\mu}) > 0$
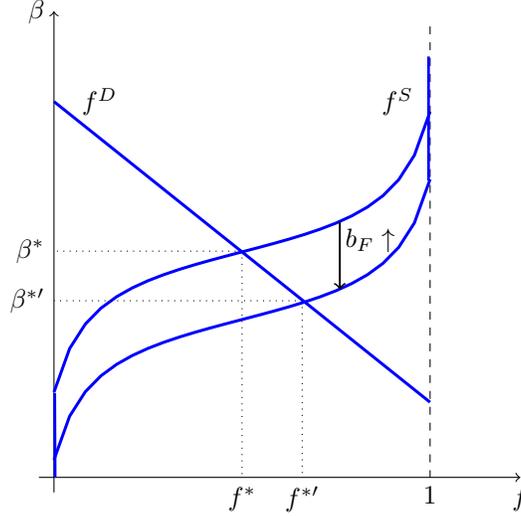
Figure 4: Increase in formal sector benefit

formality and participation. Workers entering the formal sector may welcome enhanced benefits for retirement, disability or unemployment, as these have important insurance effects. To the extent that burdens on employers can be reduced–e.g., by defined contribution schemes, financed by workers, and perhaps general revenues, the disincentive effects on firms could be minimized.

An increase in the TMIG b_O has all the same effects as the formal benefit, but in the opposite direction. On top of this, the TMIG has a direct, negative effect on participation. As we can see from equation (5), an increase in b_O will shift the supply curve up,[9] causing the formal sector wage premium β to rise and labour formality to fall as in figure 4. This comes about because even though the benefit b_O is targeted at households with no work, enforcement is imperfect and workers with informal jobs may still receive the benefit, making informal work more attractive. Since relative demand is not perfectly inelastic, the rise in the wage premium will not be sufficient to offset the increase in the consumption of informal workers coming from the increase in b_O and so informal workers will be relatively better off, making it less attractive to be a formal worker and thus decreasing labour formality.

Work in the modern sector has become less attractive overall, decreasing $\mathsf{E}\left[U\right]$. Furthermore, the TMIG b_O has a direct effect on $\tilde{\varepsilon}$ and so participation in the modern sector falls sharply, decreasing labour productivity and output. This decrease in participation reduces the supply of effective labour and so has the general equilibrium effect of increasing the cost of effective low-skilled labour $\tilde{w}_L$ which will shift the relative demand curve up, but this second order effect will not reverse the overall decrease in labour formality and especially participation.

This suggests that using targeted minimum income guarantees targeted to those outside the

---

[9]It shifts the supply curve down because $\frac{\partial f^S}{\partial b_O} = \frac{\partial f^S}{\partial \tilde{\mu}} \frac{\partial \tilde{\mu}}{\partial b_O} = -j\left(\tilde{\mu}\right) < 0$
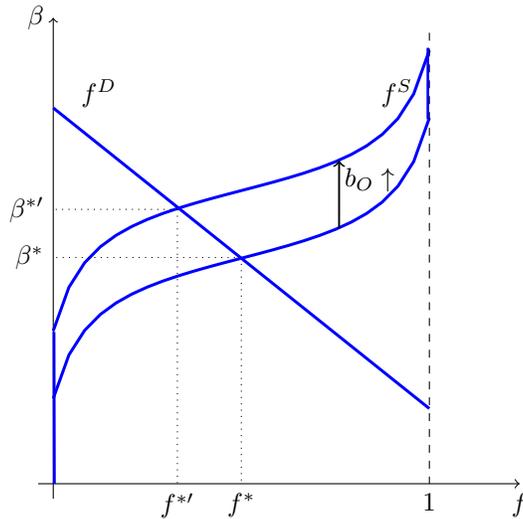
Figure 5: Increase in targeted minimum income guarantee

modern sector can be a very costly way to redistribute to this group if the level of the benefit is too high. Since, by definition, informal workers are difficult to monitor, it will be extremely difficult, if not impossible to stop informal workers from claiming the benefit. However, formal workers are registered and so it is relatively simple to prevent them from receiving a benefit they are not eligible for. This means that a high level of b_O will make it more attractive to be an informal worker, raise informality, and mean that formal employers have to raise formal wages in order to attract workers, effectively taxing formal employment, reducing productivity and efficiency.

The universal benefit b_U has the appealing feature that it does not distort incentives. Since all citizens are eligible independent of whether they are formal workers or informal workers or whether they work in the modern sector or not, the level of the universal benefit does not affect decisions to be in any of these occupations. However, this feature also makes it extremely expensive, as it must be provided for the entire population. This means that providing universal benefits may place severe demands on the government budget if they are to be provided at levels that achieve the government's social goals. Financing this may then require resorting to very high levels of taxation and the facing the resultant very strong loss of economic efficiency, productivity and growth.

Finally, we have the targeted or basic benefit b_L. A basic benefit could be useful in compensating the losers of reform, especially as a shift from the payroll tax towards a VAT would change relative prices and have an impact on middle and lower income groups. In the Mexican setting, this might mean raising the level of the Progresa/Oportunidades benefit. However, a significant portion of the informal workers are likely to be single migrants without children, the Oportunidades program may have only partial coverage. Thus, Oportunidades may need to be

supplemented by other measures.

Highly targeted versions of basic benefits (such as the BISP in Pakistan, which targets the poorest in rural areas based on an asset holding test) may be even less effective in providing compensation for those affected by the relative price changes. Many of the workers and household that would be affected by tax-based price changes are likely to be largely urban, and without access to home grown produce that may be relatively immune to the price changes.

Overall, an appropriate benefit system has to be designed with country characteristics in mind. Very specifically targeted measures may be inadequate in protecting the urban lower middle and poorer classes from the price shocks. The specific context in each country may yield different combinations of social policies or benefit systems.

In our framework, the basic benefit b_L does not affect incentives, since the characteristic on which eligibility depends (low ability) is assumed to be exogenous to any decisions by households. Of course, if households make human capital investments in order to become high ability workers, then the incentives to make this investment will be affected if the returns to the investment are reduced through the erosion of benefits. In general, the incentive effects of targeting mechanisms are an issue to be considered carefully in the design of the implementation of targeted transfers, an issue we turn to next.

4.3   Implementation

The government has two implementation parameters, γ, the effectiveness of the targeting technology with which to deliver the targeted or basic benefit b_L, and η the parameter controlling the probability of detection of sales under-reporting. These parameters are only partly exogenous and careful design can help to improve them.

The targeting parameter γ will depend on the choice of criteria used to determine eligibility for the benefit. As shown by Akerlof, 1978, good eligibility criteria have the properties of being correlated with ability, being exogenous to decisions made by households and being immutable in response to the benefit. To this we would add two issues. First in the context of many developing countries, complex eligibility criteria open up the possibility that they will be used for rent seeking purposes by officials who can manipulate eligibility scores in order to extract rents from potential recipients. Second, in a multilevel government, the issue of the appropriate level of government to finance and implement such a benefit comes to the fore. For example, in the context of India's guaranteed work programme NREGA, Niehaus & Sukhtankar (2011) show that local government (Panchayat) officials exploit changes in state level laws and federal funding to extract rents from program participants.

We touched upon the enforcement parameter η in section 3.1 but this issue bears fleshing out. Improvements in η improve the reporting of sales. This broadens the tax base of and increases revenues from both the VAT and the corporate income tax, thus raising more revenue at a lower efficiency cost. The crucial issue here is the information generated by the tax system which has two aspects. First, the tax administration authorities should optimally capture and

use the information generated by the filing of tax returns. This means effective information management systems and the sharing of information across tax bases that is relevant to both. An example of best practice in this is the recent merger of the corporate income tax and VAT departments at Her Majesty's Revenue and Customs in the United Kingdom partly in order to be able to pool information generated by each of these taxes.

The second critical component is the design of tax policy that minimises the opportunities for evasion. Exemptions from the VAT, for example, break production chains and hurt the enforcement of the VAT. However, this also has the knock-on effect of making it easier for these firms to evade their corporate income tax liabilities. In general, complicated rate structures with numerous zero-ratings and exemptions open up loopholes that firms can use to evade and avoid their tax liabilities across taxes and are therefore to be avoided. Rather than using exemptions and zero-ratings of necessities for redistributional purposes, a well targeted benefit to offset price increases on these goods, if available, is a far more efficient means of achieving redistributional aims.

5   Extensions

In order to explore how the effects we have analysed here travel along production chains, the first extension to the framework we propose would be to extend the model to include multiple sectors and intermediate inputs. For example, one of the greatest merits of a uniform VAT is that it generates powerful incentives along production chains to report transactions truthfully through the credit invoicing method of implementation. One firm may wish to evade the VAT by under-reporting its sales, but if these sales are to another firm, this would increase the VAT liability of the downstream firm and so if they were to evade the VAT they would have to collude. The longer are production chains, the less likely that this sophisticated level of collusion would be sustainable and so the VAT with invoice crediting is virtually self-enforcing. How this story changes when one or more sectors along the production chain are excluded from the VAT could be explored in a multi-sector version of our model.

A second extension we would like to explore would be to embed this framework into a multilevel government. In practice, certain tax bases are assigned to different levels of government and the revenues from different taxes are shared in different ways amongst the various levels of government. Thus, the extent that different regions differ, any reform to any of the taxes or benefits would create gainers and losers and this would be a first order concern in the passage of any proposed legislation. For instance, the Mexican case becomes quite complicated, given the overall pressures on general revenues with declining oil production, and specifically significant are constraints at the state level. This is because the nomina (payroll tax) is assigned to the states and is one of the few major own-source revenues. Complicated revenue-sharing arrangements imply that there will be gainers and losers among the states/provinces, and again compensation mechanisms through minimum intergovernmental transfers, or hold harmless con-

ditions, may be needed. A multi-government version of our framework would permit us to study these issues more carefully. Of course, the intergovernmental dimensions are more important in some countries (such as in India and Pakistan, Indonesia, as well as the proposed extensions of the empirical work to Argentina and Brazil) than others (such as in the GCC), and again, a case by case assessment is likely to be needed.

6   Conclusions and Further Work

We have presented a framework for the analysis of tax and benefit policy in contexts where informality is prevalent. We consider a wide range of taxes and benefits and highlight how they affect incentives for firms and workers to be informal and how benefit policy interacts with tax policy. Our framework allows us to consider various tax instruments jointly, in contrast to much existing work, and also consider benefit policy jointly with tax policy. We are thus able to highlight the important interactions between taxes and between taxes and benefits. We find that payroll taxes and targeted minimum income guarantees targeted at households operating outside the modern sector are particularly damaging to incentives for formality, productivity and economic efficiency. Furthermore, we find that corporate income taxes as well as the VAT have an important role to play in raising revenues in an efficient manner. We also suggest that, in addition to the traditional consideration of the efficiency cost of taxation, the impact on incentives to evade and for informality should be central considerations in the design of tax policy in developing countries.

In the coming months, we plan to use our framework as a guide for empirical work in countries around the world, including Mexico, Pakistan, Argentina, Indonesia, and possibly Gulf countries and Southern European countries to estimate the parameters that our model suggest are key to evaluating the welfare effects of reforms to the various taxes and benefits. Thus, we hope to be able to use this general framework not only to contribute estimates of key empirical entities, but to guide the development of policy proposals tailored to the particular circumstances of a wide variety of countries.

A Technical Appendix on Existence of Equilibrium

The firm's first order conditions for profit maximisation are

$$\tilde{p}\alpha_L\theta L_E^{\alpha_L-1}H^{\alpha_H} - (1-\tau_c)(1+\tau_p)w_F + F_L\left(\frac{L_I}{L}\right)^2 \leq 0 \tag{7}$$

$$\tilde{p}\alpha_L(1-\theta)L_E^{\alpha_L-1}H^{\alpha_H} - w_I - F_LL_I\left(\frac{L+L_F}{L^2}\right) \leq 0 \tag{8}$$

$$\tilde{p}\alpha_H L_E^{\alpha_L}H^{\alpha_H-1} - (1-\tau_c)(1+\tau_p)w_H \leq 0 \tag{9}$$

$$p\left\{2\eta F_Y\left(1-\frac{\hat{Y}}{Y}\right) - [1-(1-\tau_c)(1-\tau_v)]\right\} \leq 0 \tag{10}$$

where

$$\tilde{p} = \frac{\partial\Pi}{\partial Y} = p\left[1 - \eta F_Y\left(1-\left(\frac{\hat{Y}}{Y}\right)^2\right)\right] \tag{11}$$

is the firm's marginal benefit of output. Equilibrium will require that all the markets clear. That is, that the market for high-skilled labour clears:

$$H^D = H^S$$

The market for (effective) low-skilled labour clears

$$L_E^D = L_E^S$$

The submarkets for formal and informal low-skilled labour clear

$$L_F^S = L_F^D \ \& \ L_I^S = L_I^D$$

and finally that the market for final output clears:

$$
\begin{aligned}
p^*Y^* &= [w_H + b_F + b_U + (1-\gamma)b_L] + N_I(w_I + b_U + b_O + \gamma b_L) + N_F(w_F + b_U + b_F + \gamma b_L) \\
&\quad + N_O(b_O + \mathsf{E}[\varepsilon|\varepsilon < \tilde{\varepsilon}]) + \Pi^*
\end{aligned}
$$

By Walras' law we will only need to focus on 3 of these so we will focus on the first three. The first order conditions (7) - (9) imply that

$$\frac{\tilde{w}_L}{\alpha_L}L_E = \frac{\tilde{w}_H}{\alpha_H}H$$

where $\tilde{w}_H = (1 - \tau_c) (1 + \tau_p) w_H$ is the producer cost of labour, so that taking (9) we get that

$$\tilde{p}\alpha_H \left(\frac{\alpha_L}{\tilde{w}_L} \frac{\tilde{w}_H}{\alpha_H}\right)^{\alpha_L} H^{\alpha_H + \alpha_L - 1} = \tilde{w}_H$$

$$\Longleftrightarrow H^D = \left(\tilde{p} \left(\frac{\alpha_H}{\tilde{w}_H}\right)^{1-\alpha_L} \left(\frac{\alpha_L}{\tilde{w}_L}\right)^{\alpha_L}\right)^{\frac{1}{1-\alpha_L-\alpha_H}}$$

and hence that

$$L_E^D = \left(\tilde{p} \left(\frac{\alpha_H}{\tilde{w}_H}\right)^{\alpha_H} \left(\frac{\alpha_L}{\tilde{w}_L}\right)^{1-\alpha_H}\right)^{\frac{1}{1-\alpha_L-\alpha_H}}$$

which means that output is

$$Y = \left(\tilde{p}^{\alpha_L + \alpha_H} \left(\frac{\alpha_H}{\tilde{w}_H}\right)^{\alpha_H} \left(\frac{\alpha_L}{\tilde{w}_L}\right)^{\alpha_L}\right)^{\frac{1}{1-\alpha_L-\alpha_H}}$$

Labour market Equilibrium

There is a supply of 1 unit of high-skilled labour, so $H^S = 1$ meaning that the equilibrium condition for high skilled labour will require that

$$H^S = H^D$$
$$1 = \left(\tilde{p} \left(\frac{\alpha_H}{\tilde{w}_H}\right)^{1-\alpha_L} \left(\frac{\alpha_L}{\tilde{w}_L}\right)^{\alpha_L}\right)^{\frac{1}{1-\alpha_L-\alpha_H}}$$

where H^D is clearly decreasing in the wage $\tilde{w}_H$ and supply is constant so there will be a wage that clears the market.

For the low skilled, things are somewhat different as the supply of effective low-skilled labour depends on the number of formal and informal low-skilled workers.

$$L_E^S = \theta L_F^S + (1 - \theta) L_I^S$$
$$= NK(\tilde{\varepsilon}) \{\theta [1 - J(\tilde{\mu})] + (1 - \theta) J(\tilde{\mu})\}$$

so that equilibrium for effective low-skilled labour will require that

$$L_E^S = L_E^D$$
$$NK(\tilde{\varepsilon}) \{\theta [1 - J(\tilde{\mu})] + (1 - \theta) J(\tilde{\mu})\} = \left(\tilde{p} \left(\frac{\alpha_H}{\tilde{w}_H}\right)^{\alpha_H} \left(\frac{\alpha_L}{\tilde{w}_L}\right)^{1-\alpha_H}\right)^{\frac{1}{1-\alpha_L-\alpha_H}}$$

The right hand side here is decreasing in $\tilde{w}_L$ so that demand is downward sloping. Showing that the left hand side is upward sloping is a bit more involved. The relationship between

w_F and w_I, given by β is pinned down by the equilibrium in the submarkets for formal and informal labour (the subject of section 3.2). An increase in $\tilde{w}_L$ shifts the relative demand curve upwards so that both β and f are increasing in $\tilde{w}_L$. Since $L_E^S = NK\left(\tilde{\varepsilon}\right)\left[\theta f + \left(1-\theta\right)\left(1-f\right)\right]$ is increasing in f and by equation (4) $\tilde{w}_L$ is increasing in w_I, this also means that L_E^S is increasing in w_I. Since supply is increasing in $\tilde{w}_L$ and demand is decreasing in $\tilde{w}_L$ there will exist an informal wage w_I which generates a $\tilde{w}_L$ that clears the market for effective low-skilled labour.

References

AHMAD, EHTISHAM, & STERN, NICHOLAS. 1991. *The Theory and Practice of Tax Reform in Developing Countries.* Cambridge, U.K.: Cambridge University Press.

AKERLOF, GEORGE. 1978. The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. *American Economic Review*, **68**, 8–19.

ANTÓN, ARTURO, HERNÁNDEZ, FAUSTO, & LEVY, SANTIAGO. 2011. *The End of Informality in Mexico?: Fiscal Reform for Universal Social Insurance.* mimeo: Inter-American Development Bank.

ATKINSON, ANTHONY, & STIGLITZ, JOSEPH E. 1976. The Design of Tax Structure: Direct Versus Indirect Taxation. *Journal of Public Economics*, **6**, 55–75.

BOADWAY, ROBIN, & SATO, MOTOHIRO. 2009. Optimal Tax Design and Enforcement with an Informal Sector. *American Economic Journal: Economic Policy*, **1**, 1–27.

BOSCH, MARIANO, & MALONEY, WILLIAM F. 2010. Comparative analysis of labor market dynamics using Markov processes: An application to informality. *Labour Economics*, **17**, 621–631.

CHETTY, RAJ. 2009. Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics*, **1**, 16.1–16.37.

DE MEL, SURESH, MCKENZIE, DAVID, & WOODRUFF, CHRISTOPHER. 2010. *What is the Cost of Formality? Experimentally Estimating the Demand for Formalization.* mimeo, University of Warwick.

DE PAULA, ÁUREO, & SCHEINKMAN, JOSÉ A. 2010. *The Informal Sector: An Equilibrium Model and Some Empirical Evidence from Brazil.* University of Pennsylvania, PIER Working Paper 10-024.

DHARMAPALA, DHAMMIKA, SLEMROD, JOEL, & WILSON, JOHN DOUGLAS. 2011. Tax Policy and the Missing Middle: Optimal Tax Remittance with Firm-Level Administrative Costs. *Journal of Public Economics*, **95**, 1036–1047.

DIAMOND, PETER, & MIRRLEES, JAMES. 1971. Optimal Taxation and Public Production I: Production Efficiency. *American Economic Review*, **61**.

EMRAN, M. SHAHE, & STIGLITZ, JOSEPH E. 2005. On Selective Indirect Tax Reform in Developing Countries. *Journal of Public Economics*, **89**, 599–623.

EUROPEAN COMMISSION, THE. 2010. *Commission Staff Working Document accompanying the Green Paper on the Future of VAT.* SEC(2010) 1455 final, Brussels (http://ec.europa.eu/taxation_customs/resources/documents/common/consultations/tax/future_vat/sec%282010%291455_en.pdf).

GÜNTHER, ISABEL, & LAUNOV, ANDREY. 2012. Informal Employment in Developing Countries: Opportunity or Last Resort? *Journal of Development Economics*, **97**, 88–98.

KANBUR, RAVI. 2009. Conceptualizing Informality: Regulation and Enforcement. *Indian Journal of Labour Economics*, **52**, 33–42.

KEEN, MICHAEL. 2008. VAT, tariffs, and withholding: Border taxes and informality in developing countries. *Journal of Public Economics*, **92**, 1892–1906.

KEEN, MICHAEL, & MINTZ, JACK. 2004. The Optimal Threshold for a Value-Added Tax. *Journal of Public Economics*, **88**, 559–576.

LA PORTA, RAFAEL, & SHLEIFER, ANDREI. 2008. The Unofficial Economy and Economic Development. *Brookings Papers on Economic Activity*, **39**, 275–363.

LEVY, SANTIAGO. 2008. *Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico.* Washington D.C.: Brookings Institution Press.

MALONEY, WILLIAM F. 2004. Informality Revisited. *World Development*, **32**, 1159–1178.

McKENZIE, DAVID, & SAKHO, YAYE SEYNABOU. 2010. Does it Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability. *Journal of Development Economics*, **91**, 15–24.

MIRRLEES, JAMES A. 1971. An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies*, **38**, 175–208.

NIEHAUS, PAUL, & SUKHTANKAR, SANDIP. 2011. *The Marginal Rate of Corruption in Public Programs.* mimeo: University of California at San Diego.

PERRY, GUILLERMO E., MALONEY, WILLIAM F., ARIAS, OMAR S., FAJNZYLBER, PABLO, MASON, ANDREW D., & SAAVEDRA-CHANDUVI, JAIME. 2007. *Informality: Exit and Exclusion.* Washington D.C.: World Bank.

RAUCH, JAMES E. 1991. Modelling the Informal Sector Formally. *Journal of Development Economics*, **35**, 33–47.

SAEZ, EMMANUEL. 2001. Using Elasticities to Derive Optimal Tax Rates. *Review of Economic Studies*, **68**, 205–229.