Informal Sector, Government Policy and Institutions*

S. Borağan Aruoba
University of Maryland

April 4, 2010

Abstract

We document cross-country differences in informal activity, government policies and institutions using a data set covering 118 countries. Five key facts emerge: better institutions are associated with lower inflation, higher income tax rates and less informal activity and higher levels of informal activity are associated with lower income tax rates and higher inflation. We develop a general equilibrium model where households optimally choose the extent of informal activity and a benevolent government optimally chooses policies, both taking as given the institutions of the economy. The model is able to account for most of the cross-country differences in policies and informal activity as well as other key facts that emerge from the data. The performance of the model is significantly reduced for various subsets of countries, where some its key assumptions are likely to be violated.

Key Words: Ramsey problem, Friedman rule, inflation, taxation

*The author thanks Peter Murrell for extensive discussions, Stefania Albanesi, Morris Davis, David Givens, David Lagakos, Ellen McGrattan, Guillermo Ordoñez, Richard Rogerson, John Shea, Michael Waugh, Mark Wright, Randy Wright, participants at numerous conferences and seminars for useful comments and Henry Siu for collaborations in the earlier phases of this project. The usual disclaimer applies.

1 Introduction

There is considerable heterogeneity across countries regarding the sources of revenues of the government, in particular the use of income taxes versus the inflation tax. To explain this heterogeneity, we develop a model that focuses on the effect of institutions on the optimal decisions of governments via the latter's impact on the level of informal activity.[1] We demonstrate that our model is able to account for the cross-country heterogeneity in policies both qualitatively and quantitatively.

To document the cross-country differences in institutions, government policies and levels of informal activity, we compile a comprehensive data set of 118 countries. Five key facts emerge: better institutions are associated with lower inflation, higher income tax rates and less informal activity, and higher levels of informal activity are associated with lower income tax rates and higher inflation. In order to account for these facts, we use a general equilibrium model that generates both the government policies and the extent of informal activity *endogenously,* taking as given the institutional structure of the country. Our maintained hypothesis is that all countries in the world are populated by identical people and these people and benevolent governments respond optimally to economic incentives stemming from, among other factors, the institutions of the country. Considering the large set of countries we have in our dataset, the assumption of a benevolent and optimizing government choosing policies is clearly a stretch. In many countries factors other than the ones considered in our model, such as political considerations, are certainly key in determining policies. We view this as a test of our theory. We expect our theory to *fail* in such countries, and to *succeed* in those where the political considerations are of secondary importance.

There are three key components of our model. First, we explicitly model the private sector's informal activity choice. Facing a risk of a tax audit (and a punishment if found evading taxes), and taking into account the government's income tax and inflation policies, the agents in the economy optimally choose the level of informal activity.

Second, we consider institutions of the country as exogenous. In our model, these institutions determine the implicit cost of evading taxes (through the tax audit). Acemoglu,

[1] By institutions we refer to the set of rules that determine how economic activity is conducted. In our empirical analysis we use "rule of law" to measure institutions. For our purposes the terms "unofficial", "informal" or "shadow" economy refer to the same phenomenon, which is any economic activity that is done outside the reach of the government and therefore is not subject to taxation. A key characteristic of informal activity is that it is typically cash-intensive. While tax evasion is of course illegal, our concept of informal sector will not include activities that are inherently illegal.

Johnson and Robinson (2005) argue that the institutions of a country should be thought of as endogenously determined in a dynamic model, along with a number of economic and political outcomes. However, as they and many others emphasize, institutions are very persistent and thus evolve slowly – certainly much slower than the two government policies we consider. As such, one can think of our model as embedded within one period of the dynamic framework in Acemoglu, Johnson and Robinson (2005), during which institutions are constant and therefore given.

Third, we consider a benevolent and optimizing government whose objective is to raise a given amount of revenue in the least distorting way, in the tradition of Ramsey (1927). As in many similar optimal policy problems, the government strikes a balance between inflation, which is an implicit tax on cash-intensive activities and explicit taxation. The additional wrinkle in this model comes from the fact that these policies also affect the tax evasion incentives for the private sector – higher inflation deters and higher taxes encourage informal activity. Thus, the government may choose to use inflation as a tool to reduce informal activity and increase the tax base, in addition to the pure revenue motive in standard models.

To understand how the mechanism in our model works, consider two countries A and B, which are identical in all aspects except that country B has "better" institutions. The citizens of country B will choose to do less informal activity than those of country A because the cost of tax evasion is higher. This makes the marginal utility of informal activity relative to formal activity higher in country B compared to country A. Since the governments want to balance distortions, the government of country B will choose a lower level of inflation and higher level of taxes compared to that of country A. In the end, then, we find that country B has less informal activity, lower inflation and higher taxes, which shows that our model can *qualitatively* explain the five facts we obtained from the data.[2]

We then take our model to the data and conduct a cross-country exercise using the 118 countries in our sample, maintaining the assumption that abilities and desires of households are identical across countries, while three exogenous variables vary across countries: institutions, labor productivity and government expenditures. Comparing model-generated policies and private-sector behavior with those from our dataset, we conclude that our model is successful quantitatively in generating both the cross-country dispersion observed in the

[2] Although very different in model and methodology, the "tax riot" equilibrium of Bassetto and Phelan (2008), where households coordinate on underreporting their incomes, resembles the equilibrium in our model when institutions are bad and informal activity is high.

data and correlations among variables. explaining the five facts we document. Among the three exogenous variables we consider, institutions emerges as the key variable that generates much of the observed dispersion in inflation and informal activity, while the level of government expenditures is primarily responsible for explaining variation in tax rates.

In order to test the validity of our assumptions, we also consider particular subsamples of countries. We find that countries with poor institutions or low output as well as countries that are classified as "not free" show significant discrepancies between the model-generated outcomes and the data. In contrast, our model is successful in matching the data for the remaining countries. This leads us to conclude that when our key assumptions, including the existence of a benevolent and optimizing planner, are approximately correct, our model is able to explain a significant portion of the cross-country dispersion in government policies.

Our work is related to a number of different strands in the literature. There is a vast empirical literature that focuses on the causes of informal activity. Johnson, Kaufmann and Schleifer (1997), Johnson, Kaufmann and Zoido-Lobaton (1997) and Friedman et al. (2000) provide empirical results suggesting that large informal markets are typically associated with institutional factors such as excessive regulation, poor enforcement of law and corruption. These results are especially useful in establishing the link between institutions and informal activity, which is one of the key facts we exploit.

There is also a large literature on political-economy explanations of cross-country differences in government policies. One of the seminal papers, Alesina and Drazen (1991), links the delay in stabilization policies to a war of attrition between different socioeconomic groups in the country which may be affected asymmetrically by the stabilization. Until the appropriate policies are enacted, the economy follows a volatile path which is also typically associated with high inflation. Their model implies, among other things, that political polarization of a country would be associated with longer periods of instability, hence higher likelihood of high inflation.[3] Albanesi (2007) also provides a political-economy explanation for differences in inflation that relies on exogenous differences in labor productivity. In her model, the poor, those with lower labor productivity, hold more of their wealth as currency and thus are more vulnerable to inflation. If a country has a more unequal income distribution, the political bargaining process favors the rich and equilibrium inflation is higher.[4] Elgin (2009) also con-

[3] Alesina, Ardagna and Trebbi (2006) test the implications of the mechanism in Alesina and Drazen (1991) and find that the delay in stabilization is shorter when the ruling executive has more control over the legislative body of the country or when the executive has more institutional constraints.

[4] Explanations for cross-country differences in policies based on the conflict between heterogenous segments of the society have also been provided by Fernandez and Rodrik (1991), Cukierman, Edwards and

siders a real political economy model to explain different tax policies across countries where political turnover provides the exogenous variation. As we discuss above, *a priori* we expect our model to fail for countries where the above-mentioned political-economy considerations are most important. To the extent that it helps explain the facts we document for some countries, we think our theory complements the existing political-economy theories.

To complement the empirical literature on the determinants of informal activity, there has been work on economic models to formalize these links. Kuehn (2007) considers the mechanisms behind informal activity in high-income countries, building a model where agents of different abilities choose whether or not they want to become workers, managers of a firm in the formal sector or managers of a firm in the informal sector. The trade-off between the latter two exists due to the probability of getting caught and being punished. While she considers the effect of institutions on informal activity, government policies are considered to be exogenous and since this is a real model, the only government policy is taxation. Quintin (2008) shows how the degree to which financing contracts can be enforced in the formal sector (a good proxy for institutions) can affect the size of the informal sector. Ihrig and Moe (2004) build a real model with an informal labor market and investigate the effects of (exogenous) tax rates and enforcement policies on informal activity.

Our work is also linked to a small theoretical literature that shows how the extent of informal activity may lead to differences in optimal policy choices by governments. Nicolini (1998) is one of the first to show theoretically in the context of a cash-good-credit-good model that tax evasion due to informal activity is a motive for inflation under optimal policy. Yesin (2004, 2006) considers the optimal policy in the same model when the government faces (exogenous) tax collection costs and finds some success in explaining different policies for a small set of countries. The extent of informal activity (the set of goods that are formal versus informal) is assumed to be exogenously fixed in these papers. Koreshkova (2006) also models the trade-off that an optimizing planner faces between taxation (and evasion) and inflation in a cash-in-advance model with costly credit. The size of the informal sector in her model is directly linked to the assumed productivity differences across formal and informal production and as such can be considered exogenous. None of these papers consider the effect of institutions on optimal government policies through their impact on the incentives of the private sector, which is our contribution to this literature.[5]

Tabellini (1992) and Laban and Sturzenegger (1994).

[5] Ahiabu (2006) explores the trade-off between tax rates in the formal sector and audits (punishments) in the informal sector but does not conduct an optimal policy exercise.

In terms of modelling strategy, we focus on two properties of informal activity: tax evasion and cash intensiveness. To capture these features, we use a search-based monetary model combining elements from recent advances in the field such as Lagos and Wright (2005), Rocheteau and Wright (2005) and Aruoba and Chugh (2010).[6]

Our work is also related to a large literature on the macroeconomic effects of institutions. Hall and Jones (1999), one of the seminal papers in this literature, show that differences in output per worker across countries can be largely explained by differences in social infrastructure, which includes their institutions. In our empirical work we use their identification strategy and more importantly we continue their tradition of relating the differences in institutions to differences in macroeconomic outcomes – taxation and inflation policies in our case.

The paper is organized as follows. In Section 2, we provide a detailed description of the data used in our cross-country exercise and the facts that emerge. In Section 3, we present our model, show the equilibrium for given policies and the Ramsey equilibrium where policies are also endogenous. In Section 4 we present our quantitative results. Section 5 concludes. The appendix provides more details on the data and the derivations.

2 Data and Facts

In this section we document the cross-country differences in government policies and establish the facts we seek to explain using the model we develop in subsequent sections. To that end, we put together a data set that covers 118 countries. The data set includes measures of institutions, informal activity, government policies and economic indicators for 1996-2004 or a subset as dictated by data restrictions. Looking ahead to our model, since our model does not have any short-run fluctuations, we want to focus on a point in time and we take averages over this short interval to prevent any idiosyncratic events or business cycles from affecting our results. Details about the data, including a list of countries and detailed sources as well as some alternative measures we used, are available in the Appendix.

[6]In this class of models, a medium of exchange is "essential" for trade in decentralized exchange. In our model the decentralized market is narrowly defined as the informal sector but these papers take a more general view of the decentralized market.

2.1 Data Sources

Our main measure of institutions is Rule of Law as reported in World Bank Governance Matters IV, from Kaufmann, Kraay and Mastruzzi (2005). We also consider some alternatives.[7] All of these measures are very highly correlated among themselves (as the first row of panel (a) of Table 1 shows) and our conclusions are unchanged using any of the alternative measures.

There are a number of alternative estimates of the size of the informal sector that differ in terms of their methodology.[8] Since the size of the informal sector is latent by its nature, all of these methods use some observed data along with some identifying assumptions to provide an estimate. For example, the currency demand approach starts with the assumption that transactions in the informal sector use cash and that any "excess" money holdings over and above what a standard money demand regression would predict can therefore be considered as a sign of informal activity. The physical input method starts with the premise that any production, formal or informal, should use some inputs such as electricity, and as such one should be able to infer the size of the informal sector by comparing the GDP imputed using these inputs and the measured GDP. Schneider (2004) uses the DYMIMIC (dynamic multiple indicators multiple causes) method, in which a set of structural equations provide causal relationships between two sets of variables and the size of the informal sector: those identified as causes of informal activity and those identified as being affected by informal activity. For example, these equations assume that burden of taxation and burden of regulation are among the causes, while various monetary and labor market variables are among those affected by informal activity. Since the DYMIMIC method provides only a relative measure across countries, Schneider (2004) combines his relative measures with absolute measures from the currency demand approach for some selected countries to compute absolute measures for all countries. It is important to emphasize that this measure intends to capture all market-based legal production of goods and services that are deliberately concealed from public authorities either to avoid paying taxes or fees or to avoid complying with some regulation.[9] In addition to this quantitative measure, we examine some alternative qualitative measures.[10] All of

[7] The alternative institutions measures are Irregular Payments from World Economic Forum Competitiveness Report and Property Rights and Freedom from Corruption from the Heritage Foundation's Index of Economic Freedom.

[8] Schneider (2004) provides a summary of some of the major methods along with a detailed bibliography.

[9] This definition therefore excludes intrinsically illegal activities such as selling drugs and home production which is not market-based.

[10] The alternative measures of the size of the informal sector are Unreported Profits and Wages, Informal

these measures are highly correlated among themselves (as the first row of panel (b) of Table 1 shows).

Turning to government policies, measuring taxes requires some care. In our model there is no distinction between tax rates and tax revenues as fraction of national output. However, in general, computing tax rates that are conformable with assumptions in macroeconomic models is a difficult task. For example most models imply that the labor income tax creates a wedge between the real wage of a worker and his marginal product. According to this definition, social security taxes that an employer pays should be included in a measure of labor income tax along with taxes paid by the worker, even though employer-paid taxes do not affect the workers take-home pay directly. Mendoza, Razin and Tesar (1994) measure consumption, labor and capital income taxes using detailed government revenue accounts for the OECD countries. We use two recent studies that extend their methodology to more countries and/or time periods. In particular, Carey and Tchilinguirian (2000) provide updated measures for the OECD countries and IMF World Economic Outlook (2003) provides measures for a small set of non-OECD countries. Combining these two sources we get only 34 countries.[11] In order to test our model with a larger set of countries, we also consider tax revenues as a fraction of GDP as an alternative measure.

We obtain a number of macroeconomic indicators from the IMF's International Financial Statistics database, Penn World Tables (PWT 6.2) and World Bank's World Development Indicators. These include inflation, output, output per worker and government spending as a fraction of GDP, which captures the size of the government. We restrict our sample to countries with less than 20% average inflation over the period 1998-2004.[12]

One issue that needs to be addressed is whether or not official estimates of GDP include any activity that could be labeled informal.[13] In our empirical analysis, we make the assumption that the macroeconomic data that we observe reflect only formal activity and do not include any information, either as explicit measurements or as adjustments, about the informal sector. Alternatively, we could make the assumption that statistical agencies fully reflect the level of informal activity in their national account estimates. The resulting

Sector and Tax Evasion from the World Economic Forum Competitiveness Report.

[11] We combine the labor income taxes (τ^h) and consumption taxes (τ^c) to create a measure of total taxes (τ) using the formula $(1-\tau) = \left(1-\tau^h\right)/\left(1+\tau^c\right)$.

[12] This eliminates 9 countries for which we otherwise have data for. Our empirical and model-based qualitative results are unchanged if we include these countries.

[13] Bureau of Economic Analysis (2001) explains the treatment of hidden and informal activities in U.S. national accounts.

ratio of informal to formal activity from this assumption is a simple transformation of the one we use and their correlation is over 0.95. As such, all of our quantitative results will be virtually unchanged. In the absence of precise information about how the statistical agency of *each* country in our sample is able to measure or estimate informal activity, we need to make an assumption and while inconsequential, we make the assumption that measured GDP excludes informal activity.

2.2 Facts

We focus on five facts regarding informal activity, government policies and institutions. The primary objectives of this paper is to explain why these government policies differ across countries and understand these five facts. Table 1 shows all correlations mentioned below and Figure 1 shows scatter plots of the benchmark measures.

Fact 1: Institutions and inflation are negatively correlated.

Rule of law, our main measure of institutions, and inflation have a correlation coefficient of -0.43. Alternative measures for institutions yield correlations between -0.41 and -0.57. The top left panel of Figure 1 shows a scatter plot of this relationship.

Fact 2: Institutions and taxes are positively correlated.

Both tax rates and tax revenues as a percentage of GDP are positively correlated with institutions, with correlation coefficients of 0.57 and 0.48, respectively. When we use alternative measures of institutions, the correlation coefficients are between 0.40 and 0.61.

Fact 3: Institutions and the size of the informal sector are negatively correlated.

There is a strong negative relationship between institutions and the size of the informal sector, with a correlation coefficient of -0.72. Looking at alternative measures for both, we find correlation coefficients between -0.58 and -0.83.

Fact 4: Inflation and the size of the informal sector are positively correlated.

Inflation and the size of the informal sector are mildly positively correlated, with a correlation coefficient of 0.28. We compute the same correlation using the alternative measures for the size of the informal sector and the correlations are between 0.54 and 0.60.

Fact 5: Tax rates and the size of the informal sector are negatively correlated.

Both tax rates and tax revenues are negatively correlated with the size of the informal sector. The correlations are -0.50 and -0.34, respectively, and the alternative measures of informal activity yield correlations ranging from -0.20 to -0.51.

2.3 Discussion

The correlations we documented above naturally cannot be interpreted as implying causation without a structural model. In what follows, we describe a model where the relationships in Facts 1, 2 and 3 are *causal* relationships where differences in government policies and size of the informal sector are in part driven by exogenous variation in institutions across countries. In this model labor productivity and the size of the government will also be exogenous. In order to explore the plausibility of this exercise, in Tables 2 and 3 we report results from some simple regressions where we investigate the determinants of inflation, taxes and size of the informal sector. In Table 2, we explain variations in inflation and our two tax measures using the three exogenous variables we listed above. Results show that rule of law is an important determinant of government policies, even in the presence of the other two exogenous variables. For taxes this is especially important as one could argue the relationship captured in the top right panel of Figure 1 and Fact 2 is simply due to countries with better institutions having larger governments. While the latter statement is correct, columns (5) and (8) show that even after controlling for level of government expenditures, institutions are still important for understanding cross-country differences in taxes.

Even though in our structural model we consider institutions as exogenous, relying on their slow-changing nature as we explained above, in an empirical analysis such as the one in Table 2, an obvious and important issue to tackle is the endogeneity of institutions. After all, if the true causation only runs from government policies to institutions, the exercise in this paper will be pointless. To show that this is not the case we follow Hall and Jones (1999) and instrument rule of law using some geographical and historical instruments.[14] Using an alternative instrument, settler mortality that Acemoglu, Johnson and Robinson (2001) has introduced, we obtain similar results, though this instrument is available for a smaller set of countries. Columns (3), (6) and (9) in Table 2 demonstrate that there is indeed a significant relationship between instrumented rule of law and government policies.

Table 3 uses the size of the informal sector as the dependent variable. Three important conclusions emerge. First, columns (1), (3) and (5) interpret Facts 4 and 5 as causal relationships, as a number of previous papers have assumed. That is, exogenous variations in inflation or taxes are interpreted as causing differences in informal activity. However com-

[14] The instruments we use are the distance from the equator, Frankel and Romer (1999) predicted trade share, fraction of the population speaking English and the fraction of the population speaking a European language. The R^2 from the first stage is 0.60, all variables are significant at the 5% level and affect institutions positively.

paring these columns, which simply recover the correlations we reported earlier, to columns (2), (4) and (6), reveals that once we control for the countries' institutions, neither inflation nor taxes are important in explaining informal sector activity. The same conclusions would go through if we use instrumented rule of law (not reported). We view this as supporting our main hypothesis that the main exogenous variation across countries is their institutions, so that the positive relationship reported in Fact 4 is simply due to exogenous changes in institutions affecting inflation and size of the informal sector in the same direction. Similarly, the negative relationship reported in Fact 5 is due to exogenous changes in institutions affecting taxes and size of the informal sector in the opposite direction. Second, among the three exogenous variables we consider, institutions is the only one that helps explain the size of the informal sector, as columns (7) and (8) shows. Finally in column (9) we use the instrumented rule of law. While the explanatory power is reduced relative to previous columns, Rule of Law is still highly significant.

3 Model

The model is based on the structure in Rocheteau and Wright (2005), who in turn build on the setup in Lagos and Wright (2005). Time is discrete and continues forever. As it is not central to the question at hand in this paper, we abstract from any aggregate uncertainty. The economy is a closed one with no interaction with the rest of the world.[15] The economy is populated by infinitely-lived households with measure $\Lambda + 1$, where $\Lambda > 1$. In every period, a formal market meets, followed by an informal market. In the formal market (FM) all households have identical desires and abilities, supply labor to a neoclassical firm, pay labor income taxes to the government at a rate τ, consume and adjust their portfolio of assets. In this market labor and goods markets are frictionless and everyone acts as price-takers. Transactions can be completed without a need for a medium of exchange. The assumption that money is not necessary in the FM means that inflation does not have a direct impact on FM consumption, which will be key in understanding some of our quantitative results below.

In the informal market (IM), measure 1 of households would like to purchase goods and measure Λ of households are able to produce goods. We label these households as buyers and

[15] This assumption is critical in two aspects. First, it rules out government revenues from foreign trade, external borrowing or foreign aid. Second, it rules out "dollarization" where, if the inflation rate becomes too high agents in the economy can start using an alternative currency.

sellers, respectively and these types are permanent. The buyers and sellers are randomly matched in the IM where it is possible for some households to be unmatched in a given period. We assume that buyers in this market are anonymous and therefore contracts are not enforceable. As a result the sellers demand a *quid pro quo* and the buyers bring money into the IM to pay for their purchases. Once a buyer-seller pair successfully matches, they bargain over the terms of trade and the buyer pays d units of money for q units of the good. This transaction occurs outside the purview of the government and therefore the proceeds are not taxed. After the IM is complete, the buyers consume the goods they purchased in the two markets where we assume the goods are Edgeworth substitutes, i.e. the marginal utility of one good increases as more of the other good is consumed.

The buyers participate in the IM at no cost. The sellers, on the other hand, face possible audits from the government. Specifically, with probability ζ a given seller is audited. If tax-evasion is found, the government imposes a utility cost of $\mathcal{P}$. This scheme resembles the one used in Bassetto and Phelan (2008) and is a "wasteful" punishment in that no one gains from it.[16] The sellers choose whether or not they want to enter the IM understanding the audit structure. In equilibrium, due to free entry, the marginal seller will be indifferent between entering and not entering, taking into account the ex-ante cost of entering, which is $\zeta\mathcal{P}$, and the actions of other sellers which determine the probability of finding a buyer. We denote $\kappa \equiv \zeta\mathcal{P}$, and in what follows we formulate the seller's problem with κ denoting the (certain) cost of entering the IM for a particular period. We interpret κ as the difficulty of evading or avoiding taxes, or generally as the rule of law, which is our preferred measure of institutions. The two components of κ, probability of a tax audit and the punishment for evading taxes, can easily be linked to the institutions of the country. The measure of sellers that pay this cost and enter the IM is denoted by n, where $n \in (0, \Lambda)$.[17]

The government's objective is to finance a constant amount of government expenditures, denoted by $\mathcal{G}$, using revenues from income taxes in the FM, seigniorage and a one-period nominal bond. The government conducts all its activities in the FM and its budget constraint is given by

$$M_{t+1} + B_{t+1} + \tau_t p_t w_t H_t = M_t + R_{t-1} B_t + p_t \mathcal{G} \tag{1}$$

[16] One can consider a number of alternative ways of punishment. One way would be to impose a cost in terms of goods, instead of utility as we do here. In this case, however, the government can in principle use the proceeds to pay for its expenditures. This would raise the possibility to use audits to raise revenue which we choose to avoid.

[17] As a technical point, in our quantitative work, we pick a Λ where $n < \Lambda$ always holds. The determination of n depends on a free-entry condition for which this restriction is key.

where M and B denote the money and bond stocks of the government, τ is the labor income tax rate, w is the real wage rate, H is the aggregate labor supply, p is the price level and R_t is the nominal return of the bond issued in period t. We assume that bonds are book entries with no tangible proof that can be carried in to the IM. This assumption guarantees that money is the only possible asset that can be used as a medium of exchange in the IM.

In principle there are three sets of variables related to the government that one can consider endogenizing: taxation and monetary policy, expenditures and tax audits. In this paper we choose to endogenize only taxation and monetary policy. One can also make $\mathcal{G}$ a choice variable of the government (or a planner) and let it depend on the (remaining) two exogenous variables: productivity and institutions. A simple regression (not reported) shows that the latter two variables have relatively low explanatory power for the former. As such while it is a worthwhile idea, the exogenous variation in our model would not provide a good explanation of cross-country variation in government expenditures. Turning to κ (or ζ and/or $\mathcal{P}$) as an endogenous variable, it is not straightforward to model the exogenous variation that would cause κ to differ across countries without, perhaps, using a political-economy model, bringing in different exogenous variables or modelling the interaction between the different segments of the society.[18] As a result, while we believe that there is a causal link between institutions and expected punishment from tax audits, we do not model it explicitly.

Since all buyers participate in the IM, the number of successful matches are given by the matching function $\mu(n)$ with $0 \leq \mu(n) \leq \min\{n, 1\}$. Accordingly, the probability that a buyer can find a seller is given by $\alpha_b \equiv \mu(n)$ and the probability that a seller can find a buyer is given by $\alpha_s \equiv \mu(n)/n$ where α_b and α_s are taken as given by the agents. Households have utility function $u(q, x)$ where x denotes the quantity of FM consumption. We make standard assumptions on the utility function: $u_q, u_x > 0$, $u_{qq}, u_{xx} < 0$, and as mentioned above we assume $u_{qx} < 0$, which makes q and x Edgeworth substitutes.[19] The sellers operate a constant-returns-to-scale (CRS) production function in the IM given by $q = Se$, where e denotes the sellers effort in the IM and S is labor productivity, which is common across markets. In the FM, a neoclassical firm operates the same CRS production

[18] Koeppl, Monnet and Quintin (2008) provide a model of endogenous institutions where the key exogenous force is the inequality of distribution of capital across agents.

[19] That $u_{qx} \neq 0$ is necessary for technical reasons. As Aruoba and Wright (2003) show for the model in Lagos and Wright (2005), which immediately applies to the model here as well, without such an assumption a dichotomy would prevail where the IM and FM variables do not interact. From a more substantive point of view, assuming $u_{qx} > 0$ makes it clear that the goods sold in the two markets are similar goods. Our notation and derivations allow for the possibility that q and x are perfect substitutes while our numerical exercises will feature less-than-perfect substitutability.

function $Y = SH$, where H is the labor they hire in a competitive market at pre-tax real wages $w = S$. Households have linear disutility of effort in the FM and IM markets.[20] Given this we can express the utility cost of production for a seller in the IM as $c(q) = q/S$.

In what follows, we first describe the optimization problems of the buyers and the sellers in the two markets and arrive at the equilibrium which takes government policies τ and R as given. We then turn to the Ramsey problem in order to endogenize the decisions of the government.

3.1 Formal Market

We use superscripts for variables to denote the type of the agent, where B, P and N denote buyers and participating and non-participating sellers, respectively. Using $W^B(.)$ to denote the value of entering the FM and $V^B(.)$ the value of entering the IM, a buyer that enters the FM faces the problem

$$W^B(\tilde{m}_t, \tilde{b}_t) = \max_{x_t^B, h_t^B, m_t^B, b_{t+1}^B} \left\{ -Ah_t^B + V^B(m_t^B, x_t^B, b_{t+1}^B) \right\} \tag{2}$$

subject to

$$P_t x_t^B = P_t S(1-\tau_t) h_t^B + \tilde{m}_t - m_t^B + R_{t-1}\tilde{b}_t - b_{t+1}^B \tag{3}$$

$$m_t^B \geq 0 \tag{4}$$

where he chooses purchases of the FM good, his labor supply and his money and bond holdings. He experiences a disutility Ah_t^B where $A > 0$ is a parameter. He then continues to the next IM with his purchases and his money holdings. The first order conditions of this problem are given by

$$\frac{A}{S(1-\tau_t)} = V_x^B(m_t^B, x_t^B, b_{t+1}^B) \tag{5}$$

$$-\chi_t + V_m^B(m_t^B, b_{t+1}^B) \leq 0, \ = 0 \text{ if } m_t^B > 0 \tag{6}$$

$$\chi_t = V_b^B(m_t^B, x_t^B, b_{t+1}^B) \tag{7}$$

[20] The assumption that the utility function in the FM features some linearity, in our case in the disutility of labor, is key for tractability of our model. This issue is discussed in detail in Lagos and Wright (2005).

where χ_t denotes the shadow value of money or simply the multiplier on (3) and is given by

$$\chi_t = \frac{A}{P_t\left(1-\tau_t\right)S} \tag{8}$$

We also observe that the value function is linear in its arguments with slopes given by

$$W_m^B(\tilde{m}_t, \tilde{b}_t) = \chi_t \tag{9}$$

$$W_b^B(\tilde{m}_t, \tilde{b}_t) = \chi_t R_{t-1} \tag{10}$$

As Lagos and Wright (2005) argue in detail, (6) shows that when it holds with equality, the money demand of buyers does not depend on their money holdings as they entered the present FM, $\tilde{m}_t$, and if $V_m^B(.)$ is strictly monotonic, then m_t^b can be uniquely determined. This is simply a result of the linearity of the disutility of labor.

The problem of a seller who enters the FM with $\tilde{m}_t$ units of money is

$$W^S(\tilde{m}_t, \tilde{b}_t) = \max\left\{ \begin{array}{c} \max\limits_{x_t^P, h_t^P, m_t^P, b_{t+1}^P} u(0, x_t^P) - Ah_t^P + V^S(m_t^P, b_{t+1}^P) - \kappa, \\ \max\limits_{x_t^N, h_t^N, m_t^N, b_{t+1}^N} u(0, x_t^N) - Ah_t^N + \beta W^S(m_t^N, b_{t+1}^N) \end{array} \right\} \tag{11}$$

where they choose between participating in the following IM and continuing to the FM next period, and where both problems are subject to

$$P_t x_t^i = P_t S\left(1-\tau_t\right) h_t^i + \tilde{m}_t - m_t^i + R_{t-1}\tilde{b}_t - b_{t+1}^i \tag{12}$$

$$m_t^i \geq 0 \tag{13}$$

for $i = P, N$. The value function for the sellers is also linear in its arguments with the slopes given by (9) and (10).

The first order conditions for a seller who chooses to participate are

$$u_x\left(0, x_t^P\right) = \frac{A}{S\left(1-\tau_t\right)} \tag{14}$$

$$-\chi_t + V_m^S(m_t^P, b_{t+1}^P) \leq 0, \ = 0 \text{ if } m_t^P > 0 \tag{15}$$

$$\chi_t = V_b^B(m_t^P, b_{t+1}^P) \tag{16}$$

where χ_t is as defined in (8), while a non-participant seller's first order conditions are

$$\begin{aligned} u_x\left(0, x_t^N\right) &= \frac{A}{S\left(1-\tau_t\right)} && (17) \\ -\chi_t + \beta\chi_{t+1} &\leq 0, \ = 0 \text{ if } m_t^N > 0 && (18) \\ \chi_t &= \beta\chi_{t+1}R_t && (19) \end{aligned}$$

where we used (9) and (10).

We assume that there is free entry to the IM by sellers (after paying the cost κ) and this implies the free-entry condition

$$u(0, x_t^P) - Ah_t^P + V^S(m_t^P, b_{t+1}^P) - \kappa = u(0, x_t^N) - Ah_t^N + \beta W^S(m_t^N, b_{t+1}^N) \tag{20}$$

We need to obtain expressions for the IM value functions and envelope conditions to characterize the optimal choices for households, which we turn to next.

3.2 Informal Market

The value function for a buyer entering the IM is given by

$$\begin{aligned} V^B(m_t^B, x_t^B, b_{t+1}^B) &= \alpha_b\left[u(q_t^B, x_t^B) + \beta W^B(m_t^B - d_t^B, b_{t+1}^B)\right] \\ &\quad + (1-\alpha_b)\left[u\left(0, x_t^B\right) + \beta W^B(m_t^B, b_{t+1}^B)\right] \\ &= \alpha_b\left[u(q_t^B, x_t^B) - u\left(0, x_t^B\right) - \beta d_t^B\chi_{t+1}\right] \\ &\quad + u\left(0, x_t^B\right) + \beta W^B(m_t^B, b_{t+1}^B) \end{aligned} \tag{21}$$

where $\left(q_t^B, d_t^B\right)$ denotes the terms of trade the buyer faces and we used the linearity of the FM value function from (9). The first term shows that in the event the buyer is able to match with a seller, he purchases q_t^B units of the IM good, enjoys the utility of consuming this good together with the goods he bought in the FM and exits the market with d_t^B less money. The second term shows that if he cannot meet a seller he simply consumes his FM goods and proceeds to the next FM.

Similarly, the value function for a participating seller entering the IM is

$$\begin{aligned} V^S(m_t^P, b_{t+1}^P) &= \alpha_s\left[-c(q_t^S, S) + \beta W^S\left(m_t^S + d_t^S, b_{t+1}^P\right)\right] + (1-\alpha_s)\beta W^S\left(m_t^S, b_{t+1}^P\right) \\ &= \alpha_s\left[-c(q_t^S, S) + \beta d_t^S\chi_{t+1}\right] + \beta W^s\left(m_t^S, b_{t+1}^P\right) \end{aligned} \tag{22}$$

where $\left(q_t^S, d_t^S\right)$ denote the terms of trade the seller faces and linearity of the FM value function simplifies the expression. The first term shows the payoff to the seller when he meets a buyer, in which case he incurs a utility cost but acquires more money to spend in the next FM, and the second term shows that if he is not able to meet a buyer, he moves on to the next FM.

The terms of trade in the IM are determined via proportional bargaining where the buyer receives θ of the surplus and the seller receives $1-\theta$ of it. This bargaining protocol has a number of virtues over, say generalized Nash bargaining which are described in detail in Aruoba, Rocheteau and Waller (2007).[21] The outcome of the bargaining will be $d = m_t^B$, so that the buyer spends all his money, while q_t solves

$$\frac{u(q_t, x_t^B) - u\left(0, x_t^B\right) - \beta m_t^B \chi_{t+1}}{-c(q_t, S) + \beta m_t^B \chi_{t+1}} = \frac{\theta}{1-\theta} \tag{23}$$

where the numerator on the left hand side is the surplus of the buyer as shown in (21) and the denominator is the surplus of the seller from (22). This expression simplifies to

$$\beta m_t^B \chi_{t+1} = g\left(q, x_t^B, S\right) \tag{24}$$

where $g(.)$ is a combination of some primitive utility functions

$$g\left(q_t, x_t^B, S\right) \equiv \theta c(q_t, S) + (1-\theta)\left[u(q_t, x_t^B) - u\left(0, x_t^B\right)\right] \tag{25}$$

With the IM problem laid out, we are now ready to derive the relevant envelope conditions. For the buyers we get

$$V_x^B(m_t^B, x_t^b, b_{t+1}^B) = \alpha_b u_x(q_t^B, x_t^B) + (1-\alpha_b)\, u_x\left(0, x_t^B\right) \tag{26}$$

$$V_m^B(m_t^B, x_t^b, b_{t+1}^B) = \alpha_b \left[\beta \chi_{t+1} \frac{u_q(q_t^B, x_t^B)}{g_q\left(q_t, x_t^B, S\right)} - \beta \chi_{t+1}\right] + \beta \chi_{t+1} \tag{27}$$

$$V_b^B(m_t^B, x_t^b, b_{t+1}^B) = \beta W_b^B(m_t^B, b_{t+1}) = \beta \chi_{t+1} R_t \tag{28}$$

[21] The key advantage of using proportional bargaining over Nash bargaining is that the former has strong monotonicity as one of its properties, which means the payoff of the buyer strictly increases as he brings more money in to the FM. In our Ramsey problem, as Aruoba and Chugh (2010) show in a related problem, with θ sufficiently away from unity, optimal policy under Nash bargaining becomes the Friedman rule since the Ramsey planner tries to fix the inefficiency caused by the non-monotonicity of the Nash solution. In contrast, with proportional bargaining, the Friedman rule is never optimal for any θ. Given that our quantitative exercises feature positive interest rates, using proportional bargaining is a better alternative.

where in the last expression we used (10).[22] For participating sellers we get

$$V_m^S(m_t^P, b_{t+1}^P) = \beta\chi_{t+1} \tag{29}$$

$$V_b^S(m_t^P, b_{t+1}^P) = \beta W_b^S(m_t^P, b_{t+1}^P) = \beta\chi_{t+1}R_t \tag{30}$$

3.3 Household Optimality

Putting together everything we obtained so far, we can summarize our results with the following proposition.

Proposition 1 *Optimality for the households entails the following:*

a. *Participating or nonparticipating sellers will choose not to hold any money and buyers will hold money, i.e.* $m_t^B > m_t^P = m_t^N = 0$. *We denote the money holdings of buyers with* m_t.

b. *All households will hold the same quantity of bonds, which we denote by* b_t.

c. *Participating and non-participating sellers choose the same level of consumption in the FM which we denote by* x_t^S.

d. *Given the heterogeneity in the experiences of households in the previous IM, there will be 4 types of households in a given FM: matched/unmatched buyers and matched/unmatched sellers. These households will have different levels of money holdings as they enter the FM and this will be reflected in their labor supply.*

e. *Free-entry condition is given by*

$$\alpha_s\left[-c(q_t, S) + g\left(q_t, x_t^B, S\right)\right] = \kappa \tag{31}$$

where q_t *follows from the bargaining problem.*

Proof. See Appendix B.1. ■

[22] Note that in (27) the period t IM terms are multiplied by χ_{t+1} because sellers cannot use the money they acquire in this IM until the FM in $t+1$ and the outside option of buyers (relevant in the bargaining problem) is using the money in the $t+1$ FM.

3.4 Equilibrium

Combining everything obtained so far, using M_t and B_t to denote aggregate money and bond holdings and defining $\pi_{t+1} \equiv p_{t+1}/p_t$, $\mathcal{M}_t \equiv M_t/p_t$ and $\mathcal{B}_t \equiv B_{t+1}/p_t$, we can define a monetary equilibrium.

Proposition 2 *Given $R_t \geq 1$ and τ_t, a monetary equilibrium is a list of sequences $\left\{x_t^B, x_t^S, H_t, \mathcal{B}_t, \mathcal{M}_t, n_t, q_t, \pi_t\right\}$ that satisfy*

$$u_x(0, x_t^S) = \frac{A}{S(1-\tau_t)} \tag{32}$$

$$\mu(n_t)\, u_x(q_t, x_t^B) + [1-\mu(n_t)]\, u_x\left(0, x_t^B\right) = \frac{A}{S(1-\tau_t)} \tag{33}$$

$$u_x(0, x_t^S) = \frac{\beta R_t}{\pi_{t+1}} u_x(0, x_{t+1}^S) \tag{34}$$

$$\frac{1-\tau_{t+1}}{1-\tau_t} = \frac{\beta}{\pi_{t+1}} \left\{ \mu(n_t) \left[\frac{u_q(q_t, x_t^B)}{g_q\left(q_t, x_t^B, S\right)} - 1 \right] + 1 \right\} \tag{35}$$

$$\frac{A\beta \mathcal{M}_t}{S(1-\tau_t)\pi_{t+1}} = g\left(q_t, x_t^B, S\right) \tag{36}$$

$$\mu(n_t)\left[-c(q_t, S) + g\left(q_t, x_t^B, S\right)\right] = \kappa n_t \tag{37}$$

$$\mathcal{M}_t + \mathcal{B}_t + \tau_t S H_t = \frac{\mathcal{M}_{t-1} + R_{t-1}\mathcal{B}_{t-1}}{\pi_t} + \mathcal{G} \tag{38}$$

$$S H_t = x_t^B + \Lambda x_t^S + \mathcal{G} \tag{39}$$

Proof. See the Appendix. ■

While analytical proofs are difficult to obtain due to the highly nonlinear $g(.)$ function and the existence of the free-entry condition, we compute some comparative statics numerically using the calibrated values for parameters. There are three relevant exogenous variables at this stage: tax rate, inflation rate and the level of institutions. As in all monetary models, inflation acts as a tax on money holdings and as it increases buyers bring less money into the IM. This in turn reduces the amount of goods they can purchase and by reducing the payoff to the sellers, reduce the entry of sellers. Through the substitution created through the utility function, FM consumption of buyers and therefore FM output increases. An increase in the tax rates reduce the quantity consumed by the buyers and make them hold more money in order to purchase more in the IM. Sellers' payoff in the IM increases, inducing them to enter the IM. As a result FM output falls and IM output increases. An increase in κ reduce the incentives for the sellers to enter the IM by increasing the fixed cost of entering

the IM. Since the buyers are now less likely to find sellers in the IM, they reduce their money holdings and increase their consumption in FM.

3.5 Ramsey Problem

Having defined the equilibrium, which takes the policies of the government (R, τ) as given, we now turn to endogenizing these policies. We do so using the basic idea in Ramsey's (1927) original work, as further developed by Lucas and Stokey (1983) and Chari, Christiano and Kehoe (1991). The treatment of this problem is similar to that in Aruoba and Chugh (2010).

We consider the problem of a benevolent planner, the Ramsey planner, who seeks to pick the least distorting policies (R, τ) in order to finance the given government expenditures $\mathcal{G}$. We assume that the Ramsey planner is able to commit to these policies. Mechanically, the Ramsey problem then is to find policies that maximize social welfare in the resulting equilibrium. The proposition below summarizes the Ramsey problem, which is stated in Lucas and Stokey's (1983) primal form.

Proposition 3 *The Ramsey planner's problem is to choose allocations* $\left\{x_t^B, x_t^S, q_t, n_t, H_t\right\}$ *to maximize the objective function*

$$\sum_{t=0}^{\infty} \beta^t \left\{\mu\left(n_t\right)\left[u\left(q_t, x_t^B\right) - c\left(q_t, S\right)\right] - n_t\kappa + \left[1 - \mu\left(n_t\right)\right] u\left(0, x_t^B\right) + \Lambda u\left(0, x_t^S\right) - AH_t\right\} \tag{40}$$

subject to the Present-Value Implementability Constraint (PVIC)

$$\sum_{t=0}^{\infty} \beta^t \left\{u_x(0, x_t^S)\left(x_t^B + \Lambda x_t^S\right) - AH_t + \mu\left(n_t\right) g\left(q_t, x_t^B, S\right)\left[\frac{u_q(q_t, x_t^B)}{g_q\left(q_t, x_t^B, S\right)} - 1\right]\right\} = \mathcal{A}_0 \text{ (multiplier } \xi\text{)} \tag{41}$$

where $\mathcal{A}_0 \equiv u_x(0, x_0^S)\left[R_{-1}\mathcal{B}_{-1}/\pi_0 + \mathcal{M}_{-1}/\pi_0\right]$; *the resource constraint (RC)*

$$SH_t = x_t^B + \Lambda x_t^S + \mathcal{G} \text{ (multiplier } \nu\text{);} \tag{42}$$

the uniform-tax condition (UT)

$$\mu\left(n_t\right) u_x(q_t, x_t^B) + \left[1 - \mu\left(n_t\right)\right] u_x\left(0, x_t^B\right) = u_x\left(0, x_t^S\right) \text{ (multiplier } \lambda\text{);} \tag{43}$$

the free-entry condition (FE)

$$\mu\left(n_t\right)\left[-c(q_t,S)+g\left(q_t,x_t^B,S\right)\right]=\kappa n_t \text{ (multiplier } \eta\text{)}; \tag{44}$$

and the zero-lower-bound condition (ZLB) to ensure the existence of monetary equilibrium in the form of

$$\mu\left(n_t\right)\left[u_q(q_t,x_t^B)-g_q\left(q_t,x_t^B,S\right)\right]\geq 0 \text{ (multiplier } \iota\text{)} \tag{45}$$

given $\mathcal{B}_{-1}$ *and* $\mathcal{M}_{-1}$.

This problem yields allocations $\left\{x_t^B,x_t^S,q_t,n_t,H_t\right\}$ *that are associated with the optimal policies, which in turn can be obtained using*

$$\tau_t = 1-\frac{A}{Su_x(0,x_t^S)} \tag{46}$$

$$R_t = \mu\left(n_t\right)\left[\frac{u_q(q_t,x_t^B)}{g_q\left(q_t,x_t^B,S\right)}-1\right]+1 \tag{47}$$

Proof. See Appendix B.3 ■

The PVIC is a compact way of summarizing the equilibrium conditions that the Ramsey planner is subject to. Typically, the PVIC and the RC fully summarize these conditions, but in this problem we need three more conditions. First, as is standard in any monetary version of the problem, we need to make sure that the interest rate implied by the choices of the Ramsey planner is non-negative, which is necessary for the existence of monetary equilibrium. This leads to the ZLB constraint. Second, the choice of the Ramsey planner as to the number of participating buyers should be consistent with the free-entry condition, which is guaranteed by FE. Finally, the Ramsey planner is not allowed to condition the tax rate in the FM on the type of the agent or his success in the previous IM. Because of the nonseparability of preferences buyers and sellers will have different marginal utilities of consumption and in principle the Ramsey planner can exploit this. We rule this out by imposing UT.

In solving this problem, we assume that $R_{-1}\mathcal{B}_{-1}+\mathcal{M}_{-1}=0$, which means that the government does not have any net liabilities at time -1. This directly implies $\mathcal{A}_0=0$ in the PVIC. In what follows we use the following shorthands:

$$u^M\equiv u\left(q,x^B\right),\ u^U\equiv u\left(0,x^B\right),\ u^S\equiv u\left(0,x^S\right) \tag{48}$$

We also drop all arguments of remaining functions and after solving the Ramsey problem we impose steady state.

Proposition 4 *The solution to the Ramsey problem that characterizes the optimal allocations and policies* $\left(q, x^S, x^B, n, H, \xi, \iota, \eta, \lambda, \tau, R\right)$ *is given by*

$$\mu(n)\, u_x^M + [1-\mu(n)]\, u_x^N + \xi u_x^S + \xi\mu(n)\left[g_x\left(\frac{u_q^M}{g_q}-1\right)+\frac{g}{g_q^2}\left(u_{qx}^M g_q - u_q^M g_{qx}\right)\right] \tag{49}$$
$$-\frac{(1+\xi)\,A}{S}+\lambda\left\{\mu(n)\, u_{xx}^M+[1-\mu(n)]\, u_{xx}^U\right\}+\eta\mu(n)\, g_x+\iota\mu(n)\left(u_{qx}^M-g_{qx}\right)=0$$

$$(1+\xi)\left(u_x^S-\frac{A}{S}\right)+u_{xx}^S\left(\xi x^S-\frac{\lambda}{\Lambda}\right)=0 \tag{50}$$

$$(1+\xi)\, u_q^M-(1+\eta)\, c_q+(\eta-\xi)\, g_q+\xi\frac{g}{g_q^2}\left(u_{qq}^M g_q-u_q^M g_{qq}\right)+\lambda u_{xq}^M+\iota\left(u_{qq}^M-g_{qq}\right)=0 \tag{51}$$

$$u^M-u^U-(1+\eta)\, c+g\left[\xi\left(\frac{u_q^M}{g_q}-1\right)+\eta\right]+\lambda\left(u_x^M-u_x^U\right)+\iota\left(u_q^M-g_q\right)-\frac{\kappa(1+\eta)}{\mu'(n)}=0 \tag{52}$$

$$u_x^S\left(x^B+\Lambda x^S\right)-AH+\mu(n)\, g\left(\frac{u_q^M}{g_q}-1\right)=0 \tag{53}$$

$$SH=x^B+\Lambda x^S+\mathcal{G} \tag{54}$$

$$\mu(n)\, u_x^M+[1-\mu(n)]\, u_x^U=u_x^S \tag{55}$$

$$\mu(n)\,(g-c)=\kappa n \tag{56}$$

$$\iota\mu(n)\left(u_q^M-g_q\right)=0 \tag{57}$$

along with (46) and (47).

Proof. See Appendix B.4. ■

Comparative statics along with the mechanism institutions affect private and government choices are provided in the next section.

4 Quantitative Exercises

We now turn to exploring the success of our model in matching the five facts we identified in Section 2. We do this numerically, calibrating the model, since the model does not yield

unambiguous analytical results and we impose some discipline using appropriate calibration targets. To reiterate, we have three sources of exogenous variation across countries: their institutions (κ), labor productivity (S) and government spending ($\mathcal{G}$). After calibrating the model, our main goal will be to compare the implied policies and extent of informal activity from the data with their model counterparts where the countries will be assumed identical except for these three sources of exogenous variation.

4.1 Functional Forms and Calibration

We assume that buyers and sellers in the IM are matched via the urn-ball matching function given by

$$\mu(n) = n\left[1 - \exp\left(-\frac{1}{n}\right)\right] \tag{58}$$

This matching function, which has the standard constant returns to scale property, is used in labor-search models. It is obtained assuming microfoundations in which each buyer "applies" to a seller with equal probability and the probability of a given seller not finding a match is $\exp(-1/n)$.[23] We assume that households have constant-relative-risk-aversion utility over a composite good Q

$$U(q,x) = \begin{cases} \dfrac{Q(q,x)^{1-\sigma}}{1-\sigma} & \text{if } \sigma \neq 1 \\ \log\left[Q(q,x)\right] & \text{if } \sigma = 1 \end{cases} \tag{59}$$

where the composite good Q is given by the constant-elasticity-of-substitution function

$$Q(q,x) = \begin{cases} \left\{\gamma\left[(q+b)^{\varepsilon} - b^{\varepsilon}\right] + x^{\varepsilon}\right\}^{1/\varepsilon} & \text{if } \varepsilon \neq 0 \\ \left(\dfrac{q+b}{b}\right)^{\gamma} x & \text{if } \varepsilon = 0 \end{cases} \tag{60}$$

In this specification $b > 0$ is a small number to make sure $U(0,x)$ is well-defined, γ determines the relative weights of IM and FM goods and ε determines the elasticity of substitution.[24] Note that in order to preserve the Edgeworth-substitutes property of the utility function, we need $\varepsilon > 1 - \sigma$. [25] In our analysis below we use a value for ε that is less than unity, but

[23] Rogerson, Shimer and Wright (2005), page 974 provides more details.

[24] When $b = 0$, the threat points of the buyers may become undefined when, for example, $\varepsilon = 0$. For the relevant part of the domain, when $q > 0$, this utility function is virtually identical to the standard one where $b = 0$.

[25] To see this, note that $u_q = \gamma Q^{1-\varepsilon-\sigma} q^{\varepsilon-1}$ and $u_{qx} = (1-\varepsilon-\sigma)\gamma q^{\varepsilon-1} Q^{1-2\varepsilon-\sigma} x^{\varepsilon-1}$ which is negative if $1-\varepsilon-\sigma < 0$.

using $\varepsilon = 1$, i.e. assuming q and x are perfect substitutes is also admissible.[26] Also note that with $\varepsilon = 0$ and $\sigma = 1$ we get $U(q, X) = \gamma \log(q + b) - \gamma \log b + \log(X)$ which is as closely we can nest the original setup in Rocheteau and Wright (2005). Finally, we assume total government spending, which we denoted by $\mathcal{G}$ above, is a fixed fraction of total FM output

$$\mathcal{G} = GY \tag{61}$$

where G is a parameter.

The underlying assumption in our numerical exercise is that every country in the world is populated by people with identical preferences. Assuming otherwise would imply that at least a part of the cross-country differences in informal activity and policies are due to different preferences. We pick the United States as our benchmark country for calibration. We set $\beta = 0.956$ based on the real return for Aaa-rated corporate bonds in the U.S . We fix $\theta = 0.5$ (egalitarian bargaining where the surplus of the match in the IM is split equally between buyers and sellers), $\varepsilon = 0.5$, $\sigma = 1$ and $b = 0.0001$. We use $\Lambda = 4$, which is large enough that in all our experiments $n < \Lambda$ is satisfied.

This leaves five objects to be calibrated: the three exogenous variables – institutions (κ^{US}), labor productivity (S^{US}) and share of government spending in GDP (G^{US}) – and two parameters A and γ, which we jointly calibrate. The size of the informal sector relative to the formal sector is a key variable we compute for the calibration and verification of the model. This measure, which we denote by $\mathcal{R}$, is computed in the same way as the measure in Schneider (2004), but, naturally, embodies the structure of our model. It is defined as

$$\mathcal{R} \equiv \frac{\mu(n)\,\mathcal{M}}{Y} \tag{62}$$

where $\mu(n)$ is the measure of matches in the IM, $\mathcal{M}$ is the real quantity of money spent in each of these trades and Y is the output in the FM.

We use the following calibration targets

$$Y = 1,\ H = 0.3 \tag{63}$$

$$G = 0.21 \tag{64}$$

[26] All our qualitative results go through with $\varepsilon = 1$ while the model needs some adjustment, such as adding decreasing returns to scale in the IM, for obtaining quantitative results.

$$\mathcal{R} = 0.086 \tag{65}$$

$$\pi = 2.35\% \tag{66}$$

where (63) are normalizations that pin down the level of labor productivity S^{US} and the parameter A, while (64) imposes the observation from the U.S. data that the average government spending was 21% of GDP in the calibration period 1998-2004. (65) matches the size of the informal sector reported in Schneider (2004) for the U.S. and (66) uses the measured inflation rate for the U.S. in our sample. The latter two targets help determine the utility parameter γ and the level of institutions for the U.S., κ^{US}.[27] As a result, the parameters (A, γ) are calibrated as $(15.98, 0.31)$ and the exogenous variables for the U.S. are $\left(\kappa^{US}, S^{US}, G^{US}\right) = (0.38, 3.33, 0.21)$. To put the value of κ^{US} in perspective, the expected punishment of tax evasion is equivalent to the utility cost of losing about 40% of money balances. This is quite high, but it is the value the model must assume in order to achieve the low value of $\mathcal{R}$ observed for the U.S.

4.2 Main Results

We now turn to the main cross-country exercise where we vary the three exogenous variables across countries, keeping the parameters at the values calibrated for the U.S. To that end, we use the mapping summarized in (49)-(57) where the exogenous variables (κ, S, G) are mapped in to the endogenous variables $\left(q, n, x^S, x^B, H, R, \tau, \text{multipliers}\right)$ along with the definition of $\mathcal{R}$ in (62).

4.2.1 Comparative Statics

In order to show (*qualitatively)* that our model is able to capture the key facts we document, we trace the effects of changing κ, holding everything else constant, in essence taking a partial derivative. These comparative statics are reported in Figure 2. Consider a very small κ which means that the expected punishment of participating in the IM is very small. Since it is very "cheap" to do informal activity most sellers participate which lead to a large informal sector. This induces the buyers to carry a lot of real money balances. Given the level of expenditures the government needs to finance, the government finds an optimal balance between taxing

[27] Note that the income tax rate τ is not targeted in this calibration and comes out to be 19.7%. This number is smaller than the one for average tax rate in our data, 27%, but fairly close to the ratio of tax revenues to GDP, 20.5%. Given that τ can be interpreted as both measures in our model, we find this aspect of the calibration satisfactory.

income in the FM and taxing money holdings, or implicitly the IM activity. As we move from the low level of κ to higher values the sellers will have less incentives to participate in the IM due to the higher cost and n goes down. Since the buyers are now less likely to find sellers to trade with they choose to bring less money to the IM and this reduces the size of the IM. In turn, the buyers will consume more in the FM as IM and FM goods are substitutes. As a result, the social marginal cost of inflation goes up (since money balances are now lower and marginal benefit of a unit of money is higher) and the social marginal cost of taxation goes down. The planner's desire to balance distortions implies that the inflation rate is now lower and the income tax rate in the FM is now higher. This argument shows that the model is able to account for Facts 1, 2 and 3 by exogenous variations in institutions since inflation goes down, taxes go up and informal activity goes down as institutions improve. Moreover, since inflation and informal activity react in the same direction to a change in institutions, and taxes and informal activity react in opposite directions, the model can account for Facts 4 and 5 as well. As κ increases measured output in the FM may go up or down since the consumptions of buyers and sellers, the only two endogenous uses of output in the FM, move in opposite directions. Finally the last panel shows that welfare, as measured in (40) goes down as κ increases. This result may seem troublesome at first since it implies that *ceteris paribus* an improvement in institutions reduces welfare. Within the context of the model it is easy to justify this result: the IM consumption good, q, is a *good*, i.e. agents in the economy like it. As such since an increase in κ reduces the total consumption of this good (mainly by reducing n), it reduces welfare. In Section 4.3 we turn to this result argue why this does *not* mean that in reality citizens of a country are worse off relative to citizens of a country with worse institutions.

4.2.2 Cross-Country Exercise

We conduct a cross-country calibration exercise where we vary institutions and other exogenous variables across countries and investigate whether (a) our model can generate dispersions for government policies and size of the informal sector similar to those observed in the data, and (b) our model can deliver similar correlations when we focus on the five facts summarized in Section 2. We need to set the values for three exogenous variables for each country: rule of law (κ^i), share of government expenditures in output (G^i) and labor productivity (S^i) where the i superscript denotes a country-specific value.

The measure of output per worker relative to the U.S. (RLP^i) from Penn World Tables

provides a way to set S^i using

$$S^i = S^{US} \times RLP^i \tag{67}$$

We have direct measures of G^i from our dataset. Finally, since the Rule of Law measure we use from Governance Matters has an arbitrary scale, we use a simple linear transformation to convert its values into κ^i.[28]

For each of the 118 countries in our sample, we solve for the Ramsey equilibrium using the common parameters and calibrated exogenous country-specific variables. Figure 3 plots four key variables obtained from the model versus their data counterparts: inflation, taxes, the size of the informal sector and measured (formal) output. As we explained before, our model does not distinguish between income tax rate and tax revenues of the government as a fraction of GDP and we report results for both measures.[29] Each panel is set up such that clusters below (above) the 45 degree line indicate that our model produces smaller (larger) numbers than those in the data. Focusing first on the correlations, we see that our model is able to generate similar cross-country variations in these four variables as in the data, given the limited exogenous variation we had in our quantitative exercise. We also report the fraction of the cross-country dispersion in each variable (calculated as the coefficient of variation) explained by our model. Our model can explain half of the cross-country variation in inflation and it generates a slightly larger cross-country variation in taxes than in the data. About two thirds of the variation in informal activity and virtually all of the variation in output can be explained by our model. Overall, we conclude that our quantitative exercise is successful in capturing the essence of the cross-country differences in these variables. Comparing levels, however, we see that our model generates on average too much inflation, income taxes that are too low and too small informal sectors, relative to the data. We turn to the possible reason behind this result in Section 4.3, but we do not think our model's failure to match *levels* is a major obstacle to our main goal of explaining cross-country *variations* in policies and informal activity.

The main test of our hypothesis that cross-country differences in institutions can explain differences in policies and extent of informal activity observed in the data is provided in Table 4. The table has three panels. The first two panels show the correlations and the

[28] The exact details of this transformation are inconsequential for our results. We use $\kappa^i = \left(ROL^i + 2\right)\left(\kappa^{US}/3.72\right)$ where ROL^i is Rule of Law. This transformation creates a κ^i variable in the range $[0.05, 0.42]$ which contains the value for the U.S. and includes sufficient variation.

[29] Since we have only 34 countries with tax data, all the subsequent results using tax rates will have a significantly smaller subsample relative to other results.

explained variation for the five key variables as reported in Figure 3. The last panel shows the correlations underlying our five facts as computed from the model-generated data. For Facts 2 and 5, which are related to taxes, we report two sets of numbers; (a) refers to those computed using tax rates and (b) refers to those computed using tax revenues.

We report results for five assumptions regarding exogenous variable as we now explain. First, our benchmark results with all three exogenous variables differing across countries show that the model-generated data replicates the correlations we reported in Section 2. That is, in addition to generating a large fraction of the cross-country variation in policies and informal activity, as we discussed above, our model can also generate the five key facts related to correlations among these variables. We should note that for Fact 1, the negative relationship between inflation and institutions, the model delivers a somewhat stronger relationship than what is in the data, possibly indicating that there are other reasons (other than the three exogenous variables we consider) why inflation rates differ across countries.[30] Second, we show the results from a cross-country exercise, in which the level of institutions is kept fixed at its U.S. value, while the other two exogenous variable vary across countries. Focusing on the correlations in panel (a) we see that the model fails to generate inflation and informal activity in line with the actual data, while the predictions for taxes and output are as good as the benchmark results. As a result of the former finding, panel (c) shows that the model cannot match Fact 1 and Fact 4, the positive relationship between inflation and the size of the informal market, which now have opposite signs of the benchmark results. Third, we shut down all exogenous variation except for institutions, analogous to the qualitative exercise we conducted in Figure 2. We find that the model with only institutions is able to generate similar variations in inflation and size of the informal sector to the data. The model is less accurate for taxes and fails completely for output. Panel (c) indicated that this model gets the signs of the correlations right, but fails to match the magnitudes completely. Fourth, to investigate the effect of labor productivity, we set the labor productivity of all countries equal to the value for the U.S. The model matches the data as well as with all three exogenous variables in all respects with the clear exception of output. Finally, we turn off both labor productivity and institutions to investigate the contribution of government spending. The model fails to generate inflation and output and to some degree the extent of informal activity in line with the data while delivering almost identical taxes to the full

[30] This is also evident in Table 2 where the R^2 of the simple regression in column (2) that explains inflation using the three exogenous variables is only 0.20.

model. Accordingly, it shows wrong signs for Facts 1 and 4.

To sum up, our analysis reveals that exogenous variations in institutions is key for understanding both the cross-country variations in informal activity and government policies as well as the five facts we documented in Section 2, while variations in labor productivity, perhaps not surprisingly, are key for cross-country variations in output but do not help explain the remaining variables. The size of the government also plays a key role in understanding cross-country variations in taxes.

4.3 Discussion and Variations

Table 5 reports results for various subsamples of countries with all three exogenous variables changing across countries.[31] First, we consider only the countries whose institutions (as measured by Rule of Law) are in the lowest quintile. The model-generated size of the informal sector is essentially uncorrelated with that in the data and the correlation for inflation is weak. As a result, the model is unable to match Fact 2, the positive relationship between taxes and institutions, Fact 4 and Fact 5, the negative relationship between size of the informal market and taxes. Next we consider countries whose outputs are less than 20% of the U.S., which is a fairly large sample of 66 countries. The correlation of actual and model-based inflation is about half and that of the size of the informal sector is less than a third of the full sample of countries. There are only four countries in this subsample for which we have tax data and the model fails to match their tax rates. The correlations in panel (c) are either much weaker than their counterparts in the full sample or, as in the case of Fact 4, have the wrong sign. The failure of our model for countries with poor institutions may be due to the fact that tax evasion is too "cheap" in reality in these countries, much more so than implied by our model. Similarly, for low output countries, our implicit assumption of perfect credit markets in the FM may be violated and this may create a divergence between our model and the data. In contrast, when we focus on the set of countries whose institutions are in the second or higher quintiles and whose outputs are at least 20% of that of the U.S., a sample of 49 countries, the results are fairly similar, if not stronger for some statistics, to those in the full sample. These results show that our model, along with all its assumptions, fail for the poor and/or low-institution countries and that its successes in the full sample are mainly driven by the remaining countries.

[31] In this table we recompute all the data-based correlations for each fact as we change samples since for some subsamples the signs of the correlations and/or their strengths change.

In Table 6 we slice countries along the "freedom" margin. We use the 2003 ratings by Freedom House, who identify "Free" and "Not Free" countries. Restricting our sample to the former group, we see that the results are virtually identical to those we obtained in the full sample. When we focus on the latter group instead, the fit between the data and the model significantly weakens, and the sign of the correlation for Fact 4 is reversed. It is also interesting to note that the Fact 2 and Fact 5 are no longer valid in the data for this set of countries and as such the fact that the model delivers signs similar to the benchmark results become counterfactual. Intuitively, this failure of our model is a clear indication that the assumption of a benevolent optimizing government choosing policies is not appropriate for these countries.

All in all, we find that our model is the most successful for countries that have sufficiently good institutions, are not very poor and are not governed by repressive or authoritarian governments (as inferred by the "not free" ranking in the Freedom House rankings). We consider it a success that our model is able to generate results in line with the data for these countries while it generates one or more counterfactual results for countries outside this group, for which many other considerations, including the political-economy ones we summarized in the Introduction, may be much more important.

Let us turn to why our model fails to match the levels of government policies and extent of informal activity. This is the immediate result of our assumption that inflation has no direct impact on FM activity, or more specifically our assumption that transactions in the FM do not require a medium of exchange. This is certainly a simplifying assumption as in any economy some of the purchases in the formal market is done using money. It is an especially critical assumption in countries where many consumers are hand-to-mouth consumers with no ability to save or are without access to credit markets. If one were to model this explicitly, FM consumption would also fall as inflation increases, adding to the cost of inflation. In our model, then, the Ramsey planner does not account for the cost of inflation on welfare through the FM and as a result chooses a level of inflation that is higher than in the data. Since inflation is too high (relative to data), informal activity is discouraged too much in the model. This also means the income tax rate is too low since sufficient revenue is raised through seigniorage. In the context of our model, we can fix this problem by, for example, splitting the FM into two parts: one where money is essential, just as it is in the IM, and one where it is not. We believe doing so, and using a measure of velocity of money in our calibrations for each country, would bring the levels of these three variables closer to those

in the data *without* affecting the cross-country variations and correlations. Since our main goal was explaining the cross-country variations, we don't pursue this idea here.

Before we conclude, we want to return to the discussion of welfare that we started in Section 4.2.1. Figure 2 showed that ceteris paribus an increase in κ leads to a decrease in welfare. However, as Hall and Jones (1999) and others have shown, changing the institutions of a country without affecting, among other things, its labor productivity is not reasonable.[32] Therefore in order to evaluate how institutions affect welfare, we need to consider also the variations in labor productivity. Figure 4 shows the welfare of each country in our dataset calculated in our cross-country exercise using (40). It is plotted versus institutions but all three exogenous variables are allowed to change across countries. This figure clearly demonstrates that welfare is an increasing function of institutions once the effect of institutions on labor productivity is accounted for.[33] When we look deeper into the results, we find, not surprisingly, that labor productivity is the main driving force of welfare in our model.

5 Conclusion

We present five key facts regarding informal activity, government policies and institutions using a data set of 118 countries. In addition to documenting that there is significant variation across countries in these dimensions, we identify five regularities. Good institutions are associated with lower inflation, higher taxes and smaller informal sectors. As a result, lower taxes and higher inflation are associated with larger informal sectors. We interpret the first set of facts as causal relationships and provide a model that delivers such relationships. The key assumption in our model is the existence of a benevolent and optimizing government, choosing tax and inflation rates, taking as given the institutions of the country and taking into account its impact on the private sector's choices of informality. The model is also successful quantitatively: it explains about half of the cross-country variation in inflation, virtually all of the variation in taxes and about two thirds of the variation in informal activity. It also delivers the five facts listed above. As a test of our model and assumptions, we then focus on some subsets of countries where one or more of our key assumptions are likely to be violated. We find that for each of these subsets, the model fails to explain some aspects of

[32] This can also be verified in our dataset. A simple IV regression of labor productivity on instrumented rule of law has an R^2 of 0.53.

[33] Perhaps to capture the precise effect of institutions on labor productivity, it is more appropriate to feed not the actual labor productivity of the country but the fitted value from the IV regression mentioned in the previous footnote. Given the high R^2 of that regression, results will be similar.

the data. For the remaining countries, our results are still valid and even stronger in some cases.

The policy takeaway from our exercise is that reforming institutions of a country, specifically reducing incentives for tax evasion, is key for increasing its output and welfare. This in itself is not a novel result in the context of the bigger literature on institutions. But what our analysis adds is to show that as the country fights (successfully) tax evasion, inflation will fall, taxes will increase, the informal sector of the economy will shrink and more of the economic activity will be registered in the formal sector. While we do not explicitly model this, labor productivity of the country will also improve, leading the an improvement in the well-being of its citizens.

Our goal in this paper was to explain the cross-country variation in informal activity and policies at a point in time. An equally interesting exercise would be to consider a dynamic framework where these variables, as well as institutions, vary over time. Our model also abstracted from capital accumulation and assumed simple (identical) linear production functions in the two sectors. A stylized fact (at least a common modeling choice among related papers) is the difference in productivity and/or factor intensities between formal and informal sectors. Extending our model to capture this fact would be an interesting exercise. Finally, a number of "stylized facts" related to inflation such as its volatility or cyclicality differ across developed and developing countries, which is a challenge for our standard models. We can address this in a dynamic and stochastic version of our model.

References

[1] Acemoglu, Daron, Simon Johnson and James A. Robinson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 1369-1401.

[2] Acemoglu, Daron, Simon Johnson and James A. Robinson (2005), "Institutions as the Fundamental Cause of Long-Run Growth," *Handbook of Economic Growth,* Philippe Aghion and Stephen Durlauf (eds), Elsevier, North Holland.

[3] Ahiabu, Stephen Eli (2006), "Optimal Audits with Search and Tax Evasion," *mimeo*, University of Toronto.

[4] Albanesi, Stefania (2007), "Inflation and Inequality," *Journal of Monetary Economics,* 54(4), 1088-1114.

[5] Alesina, Alberto, Silvia Ardagna and Francesco Trebbi (2006), "Who Adjusts and When? The Political Economy of Reforms," *IMF Staff Papers*, 53, 1-29

[6] Alesina, Alberto. and Allan Drazen (1991), "Why are stabilizations delayed?," *American Economic Review,* 81 (5), 1170–1188.

[7] Aruoba, S. Boragan, Sanjay K. Chugh (2010), "Optimal Fiscal and Monetary Policy when Money is Essential," *Journal of Economic Theory*, forthcoming.

[8] Aruoba, S. Boragan, Guillaume Rocheteau, Christopher Waller (2007), "Bargaining and the value of money," *Journal of Monetary Economics,* 54, 2636–2655.

[9] Aruoba, S. Boragan and Randall Wright (2003), "Search, Money and Capital: A Neoclassical Dichotomy", *Journal of Money, Credit and Banking*, 35, 1086-1105.

[10] Bassetto, Marco and Christopher Phelan (2008), "Tax Riots," *Review of Economic Studies*, 75, 649-669.

[11] Bureau of Economic Analysis (2001), "Hidden and Informal Economic Activities in the U.S. National Income and Product Accounts", *mimeo.*

[12] Carey, David and Harry Tchilinguirian (2000), "Average Effective Tax Rates on Capital, Labour and Consumption," *OECD Economics Department Working Papers,* No. 258.

[13] Chari, V.V., Lawrence Christiano and Patrick Kehoe (1991), "Optimal Fiscal and Monetary Policy: Some Recent Results," *Journal of Money Credit and Banking*, 23(3), 519-539.

[14] Cukierman, Alex, Sebastian Edwards and Guido Tabellini (1992), "Seignorage and Political Instability," *American Economic Review,* 82(3), 537-555.

[15] Elgin, Ceyhun (2009), "Political Turnover, Taxes and the Shadow Economy," *mimeo*, University of Minnesota.

[16] Fernandez, Raquel and Dani Rodrik, (1991), "Resistance to Reform: Status Quo Bias in the Presence of Individual- Specific Uncertainty," *American Economic Review*, 81(5), 1146-1155.

[17] Frankel, J. and D. Romer (1999), "Does Trade Cause Growth?," *American Economic Review*, 89(3), pp. 379-399.

[18] Friedman, Eric, Simon Johnson, Daniel Kaufmann, Pablo Zoido-Lobaton (2000), "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries," *Journal of Public Economics*, 76, 459–493.

[19] Hall, Robert E. and Charles I. Jones, (1999), "Why Do Some Countries Produce So Much More Output Per Worker Than Others?", *Quarterly Journal of Economics*, 114(1), pp. 83-116.

[20] Ihrig, Jane and Karine S. Moe (2004), "Lurking in the Shadows: The Informal Sector and Government Policy," *Journal of Development Economics,* 73, 541-557.

[21] International Monetary Fund World Economic Outlook (2003).

[22] Johnson, Simon, Daniel Kaufmann, Andrei Shleifer (1997), "The Unofficial Economy in Transition," *Brookings Papers on Economic Activity*, 1997(2), 159-239.

[23] Johnson, Simon, Daniel Kaufmann, Pablo Zoido-Lobatón (1997), "Regulatory Discretion and the Unofficial Economy," *American Economic Review Papers and Proceedings*, 88(2), 387-392.

[24] Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi (2005), "Governance Matters IV: Governance Indicators for 1996-2004," *World Bank Policy Research Working Paper,* No. 3630.

[25] Koeppl, Thorsten, Cyril Monnet and Erwan Quintin (2008), "Efficient Institutions," *mimeo*, Queen's University.

[26] Koreshkova, Tatyana A. (2006), "A quantitative analysis of inflation as a tax on the underground economy," *Journal of Monetary Economics,* 53, 773–796.

[27] Kuehn, Zoe (2007), "Tax Rates, Governance, and the Informal Economy in High-Income Countries", *mimeo*, Universidad Carlos III de Madrid.

[28] Laban, Raul and Federico Sturzenegger (1994), "Distributional Conflict, Financial Adaptation and Delayed Stabilizations," Economics and Politics, 6(3), 257-276.

[29] Lagos, Ricardo and Randall Wright (2005), "A Unified Framework for Monetary Theory and Policy Evaluation", *Journal of Political Economy,* 113, 463-484.

[30] Lucas, Robert E. Jr and Nancy Stokey (1983), "Optimal Fiscal and Monetary Policy in an Economy Without Capital," *Journal of Monetary Economics*, 12, 55-93.

[31] Mendoza, Enrique G., Assaf Razin, Linda L. Tesar (1994), "Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption," *Journal of Monetary Economics,* 34, 297-323.

[32] Nicolini, Juan Pablo (1998), "Tax Evasion and the Optimal Inflation Tax," *Journal of Development Economics*, 55, 215-232.

[33] Quintin, Erwan (2008), "Contract Enforcement and the Size of the Informal Economy," *Economic Theory*, 37, 395-416.

[34] Ramsey, Frank P. (1927), "A Contribution to the Theory of Taxation," *Economic Journal*, 37(145), 47-61.

[35] Rocheteau, Guillaume and Randall Wright (2005), "Money in Search Equilibrium, in Competitive Equilibrium and in Competitive Search Equilibrium", *Econometrica,* 73, 75-202.

[36] Rogerson, Richard, Robert Shimer and Randall Wright (2005), "Search-Theoretic Models of the Labor Market: A Survey", *Journal of Economic Literature,* XLIII, 959-988.

[37] Schneider, Friedrich, (2004), "The Size of the Shadow Economies of 145 Countries all over the World: First Results over the Period 1999 to 2003," *IZA Discussion Paper*, No. 1431.

[38] Yesin, Pinar (2004), "Financing Government Expenditures Optimally," *mimeo*, Study Center Gerzensee.

[39] Yesin, Pinar (2006), "Tax Collection Costs, Tax Evasion and Optimal Interest Rates," *mimeo*, Study Center Gerzensee.

Appendix

A Data

Our dataset consists of 118 countries. A list of these countries along with the values for the five key variables used in the analysis are provided on Table 7. Below are detailed information for each of the variables used in this paper.[34]

Macroeconomic Variables

- **Output** : Output relative to the U.S. from PWT (Y). Averaged over 1998-2004.
- **Output per worker** : Real GDP per worker from PWT (RGDPWOK). Used as relative to the U.S. Averaged over 1998-2004.
- **Inflation** : Annual change in CPI. From IFS (XZF). Restricted to countries with less than 20% annual inflation. Averaged over 1998-2004.
- **Government Spending** : Government Expenses (as share of GDP) from WDI. For 19 countries this series is unavailable and government's share of real GDP per capita from PWT (KG) is used instead. Averaged over 1998-2004.
- **Tax Rate** : Available for 34 countries. (See Table 7) Calculated as $\left(\tau^c + \tau^h\right) / \left(1 + \tau^c\right)$ where τ^c and τ^h are consumption and labor-income taxes, respectively. For Chile, Costa Rica, Mexico, Peru, Philippines, South Africa, Sri Lanka, Thailand and Tunisia, tax rate data comes from IMF World Economic Outlook (2003) covering unspecified periods (possibly as large as 1990-2002) for each country. For the remaining 25 countries, tax rate data come from Carey and Tchilinguirian (2000), averaged over 1991-1997.
- **Tax Revenue** : Tax Revenues (as percentage of GDP) from WDI. Available for 97 countries. (See Table 7) Averaged over 1998-2004.

Informal Activity

- **Size of Informal Sector** (Main measure): As a fraction of formal (measured) GDP. From Schneider (2004). Averaged over 1999-2003.

[34] The acronyms used are: PWT (Penn World Tables version 6.2), IFS (International Financial Statistics), WDI (World Development Indicators), WEF (World Economic Forum Global Competitiveness Report, 2001-2002). Expressions in parantheses following data sources are the data mnemonics from the original source, where available. In the case of missing data for a given country, averaging is done over the available sample.

- **Unreported Profits and Wages** : From WEF. Available for 65 countries. Answer to question 6.13 : "What amount of profits and wages does a company in your industry typically 'keep of the books'? (1=less than 5%, 2=5-10%, 3=11-20%, 4=21-30%,..., 9=71-80%, 10=more than 80%)"

- **Informal Sector** : From WEF. Available for 65 countries. Answer to question 6.14: "What percentage of businesses in your country would you guess are unofficial or not registered? (1=less than 5%, 2=6-10%, 3=11-20%, 4=21-30%,..., 8=61-70%,9=more than 70%)"

- **Tax Evasion :** From WEF. Available for 65 countries. Answer to question 6.11: "Tax evasion of your country is (1=rampant, 7=minimal)".

Institutions

- **Rule of Law** (Main measure): From World Bank Governance Matters IV. Calculated from 24 primary sources, that include a total of 74 different concepts. Sample concepts: losses and costs of crime, enforceability of government contracts, kidnapping of foreigners, organized crime, quality of police, money laundaring, property rights, independence of judiciary, fairness of the court system. Averaged over 1996-2004.

- **Irregular Payments :** From WEF. Available for 65 countries. Average of answers to questions 7.01 through 7.05 which ask: "How commonly do firms in your industry give irregular extra payments or bribes connected with import and export permits / when getting connected to public utilities / connected with annual tax payments / connected with public contracts, investment projects / connected with loan applications? (1=common, 7=never)"

- **Property Rights** : From WEF. Available for 65 countries. Answer to question 6.02: "Financial assets and wealth are (1=poorly delineated and not protected by law, 7=clearly delineated and protected by law)."

- **Freedom From Corruption :** From Heritage Foundation Index of Economic Freedom 2008. Available for 112 countries. The score for this component is derived primarily from Transparency International's Corruption Perceptions Index, which measures the level of corruption in 180 countries.

Freedom : From Freedom House, Freedom in the World 2003. Available for 117 countries. Ranks countries according to subcategories: electoral process, political pluralism and participation, functioning of government, freedom of expression and belief, associational and organizational rights, rule of law, personal autonomy and individual rights and groups them in categories: free, partly free and not free.

Instruments for Institutions

See Hall and Jones (1999) for detailed description of these variables.

- Distance from the Equator
- Log predicted trade share based on a gravity model of international trade that uses only the country's population and geographical factors, constructed by Frankel and Romer (1999).
- Fraction of the population that speaks English
- Fraction of the population that speaks a European language

B Proofs of Propositions and Details of the Model

B.1 Proof of Proposition 1

Combining (5)-(7) with (26)-(28) we get the following optimality conditions for buyers

$$\frac{A}{S\left(1-\tau_t\right)} = \alpha_b u_x(q_t^B, x_t^B) + \left(1-\alpha_b\right) u_x\left(0, x_t^B\right) \tag{68}$$

$$\chi_t \geq \beta\chi_{t+1}\left[\alpha_b\left(\frac{u_q(q_t^B, x_t^B)}{g_q\left(q_t^B, x_t^B, S\right)} - 1\right) + 1\right], \ = 0 \text{ if } m_t^b > 0 \tag{69}$$

$$\chi_t = \beta\chi_{t+1}R_t \tag{70}$$

Combining (14)-(16) with (29)-(30) we get the following optimality conditions for participating sellers

$$u_x\left(0, x_t^P\right) = \frac{A}{S\left(1-\tau_t\right)} \tag{71}$$

$$\chi_t \geq \beta\chi_{t+1}, \ = 0 \text{ if } m_t^P > 0 \tag{72}$$

$$\chi_t = \beta\chi_{t+1}R_t \tag{73}$$

Finally (17)-(19) are the optimality conditions for nonparticipating sellers, which are repeated below

$$u_x\left(0, x_t^N\right) = \frac{A}{S\left(1-\tau_t\right)} \tag{74}$$

$$\chi_t \geq \beta\chi_{t+1}, \ = 0 \text{ if } m_t^N > 0 \tag{75}$$

$$\chi_t = \beta\chi_{t+1}R_t \tag{76}$$

In all these equations we use the definition

$$\chi_t \equiv \frac{A}{P_t\left(1-\tau_t\right)S} \tag{77}$$

which is the shadow value of money in period t.

Proof of (a): In (69) the sign of the term in the square brackets, which is the return on holding money in the period t IM, depends implicitly on m_t^b. Given both (69) and (70), the buyers choose a level of $m_t^b > 0$ such that this term exactly equals R_t, which can be guaranteed as long as R_t is not too large. Intuitively, since there is no uncertainty, the return on holding money and bonds need to be identical. If, however, R_t is large and no value of m_t^b can deliver the same return, the buyer will choose $m_t^b = 0$. Comparing (69), (72) and (75), we can conclude that if $m_t^b > 0$ then $m_t^P = m_t^N = 0$ since money cannot provide the same return to the sellers (participating or non-participating) as the bond.

Proof of (b): Since (70), (73) and (76) are identical, all three types of agents choose the same level of bond holdings.

Proof of (c): Since (71) and (74) are identical, $x_t^P = x_t^N$.

Proof of (d): Given the structure of the environment and the properties of equilibrium we found so far, there are four types of agents at the start of period t, that differ according to how much money they enter period t with and how much money they exit the FM in period t.

The table below summarizes the properties and actions of these four types of agents. The MB agents are buyers who were matched in period $t-1$. Looking ahead to equilibrium, this means they will have no money at the begining of period t and will exit the FM with $m_t^B = m_t = M_t$ units of money.[35] The last column shows the measure of these agents. The UB agents are similar except they have kept their money from period $t-1$ since they could

[35] Strictly speaking, we need to use d_t^M *and* d_t^S in the discussion that follows. For clarity we impose the equilibrium outcome that $d_t^M = d_t^S = M_t$.

not find a match. The MS agents are sellers who participated and found a match in the previous IM. They acquire $m_{t-1}^B = M_{t-1}$ units of money from the buyer they met and as soon as they can, in the period t FM, they spend it. Finaly, the US agents either chose not to participate in the period $t-1$ IM, or they participated and were not matched. It is important to note that the choices of participation of MS and US agents in period t does not affect any of this.

Type	Matched in $t-1$	Type in t	$\tilde{m}_t$	m_t	Measure
MB	Yes	B	0	M_t	$\mu(n_{t-1})$
UB	No	B	M_{t-1}	M_t	$[1-\mu(n_{t-1})]$
MS	Yes	P or N	M_{t-1}	0	$\mu(n_{t-1})$
US	No	P or N	0	0	$[\Lambda-\mu(n_{t-1})]$

Denoting h_t^i, $i = MB$, UB, MS or US the labor supply choices of each type, their FM budget constraints can be written as

$$\begin{aligned}
x_t^B &= \Omega_t + S(1-\tau_t)h_t^{MB} + \left(\frac{0-M_t}{p_t}\right) \\
x_t^B &= \Omega_t + S(1-\tau_t)h_t^{UB} + \left(\frac{M_{t-1}-M_t}{p_t}\right) \\
x_t^S &= \Omega_t + S(1-\tau_t)h_t^{MS} + \frac{M_{t-1}}{p_t} \\
x_t^S &= \Omega_t + S(1-\tau_t)h_t^{US}
\end{aligned}$$

where we use $\Omega_t \equiv (R_{t-1}B_t - B_{t+1})/p_t$. Solving each of these equations for h yields

$$h_t^i = \begin{cases}
\dfrac{1}{S(1-\tau_t)}\left(x_t^B - \Omega_t + \dfrac{M_t}{p_t}\right) & i = MB \\
\dfrac{1}{S(1-\tau_t)}\left(x_t^B - \Omega_t + \dfrac{M_t - M_{t-1}}{p_t}\right) & i = UB \\
\dfrac{1}{S(1-\tau_t)}\left(x_t^S - \Omega_t - \dfrac{M_{t-1}}{p_t}\right) & i = MS \\
\dfrac{1}{S(1-\tau_t)}\left(x_t^S - \Omega_t\right) & i = US
\end{cases}$$

This shows the well-known result of the Lagos-Wright model where the agents' choices are heterogenous in the variable which enters utility linearly.

Proof of (e): Turning to the free-entry condition, the expression in (20) simplifies to

$$-Ah_t^P + V^S(0, b_{t+1}) - \kappa = -Ah_t^N + \beta W^S(0, b_{t+1})$$

since $m_t^P = m_t^N = 0$, $b_t^P = b_t^N = b_t$ and $x_t^P = x_t^N$ from (a)-(c) above. From (d), we see that $h_t^P = h_t^N$. Moreover, from (22) we get

$$V^S(0, b_{t+1}) = \alpha_s \left[-c(q_t^S, S) + \beta d_t^S \chi_{t+1}\right] + \beta W^s(0, b_{t+1})$$

Finally, equilibrium implies $d_t^S = m_t^B = M_t$. Combining these results, the free-entry condition simplifies to

$$\alpha_s \left[-c(q_t^S, S) + \beta M_t \chi_{t+1}\right] = \kappa$$

Using the result in (24), we obtain the expression in the main text.

B.2 Details of Definition 2

Most of the equilibrium conditions directly follow from the discussion above. In particular, (32) follows from (71) and (74); (33) follows from (68); (34) follows from the definition of χ_t in (77) along with (32) and (70), (73) and (76); (35) follows from (69) with equality since we are characterizing a monetary equilibrium; (36) follows from the definition of χ_t in (77) along with the outcome of the bargaining problem (24); (37) follows from the free-entry condition (31) and (38) follows from the budget constraint of the government (1). In order to obtain the resource constraint in (39) we need to define the aggregate labor supply H_t by summing up the expressions of h_t^i for each type $i = MB$, UB, MS and US using their appropriate weights. We obtain

$$\begin{aligned} H_t &= \mu(n_{t-1}) h_t^{MB} + [1 - \mu(n_{t-1})] h_t^{UB} + \mu(n_{t-1}) h_t^{MS} + [\Lambda - \mu(n_{t-1})] h_t^{US} \\ &= \frac{1}{S(1-\tau_t)} \left[x_t^B + \Lambda x_t^S + \frac{B_{t+1} - R_{t-1} B_t}{p_t} + \frac{M_t - M_{t-1}}{p_t} \right] \end{aligned} \tag{78}$$

which can also be viewed as the aggregate budget constraint of the households. Combining this with the government's budget constraint, we get (39).

B.3 Proof of Proposition 3

Social welfare function in a period can be defined as the sum of the utility functions for the four type of agents as they enter the period.

$$\begin{aligned}\mathcal{U}_t &= \mu(n_t)\left[u\left(q_t, x_t^B\right) - Ah_t^{MB}\right] + \left[1-\mu(n_t)\right]\left[u\left(0, x_t^B\right) - Ah_t^{UB}\right] \\ &\quad +\mu(n_t)\left[-c(q_t,S) + u\left(0, x_t^S\right) - Ah_t^{MS}\right] + \left[\Lambda - \mu(n_t)\right]\left[u\left(0, x_t^S\right) - Ah_t^{US}\right] - n_t\kappa \\ &= \mu(n_t)\left[u\left(q_t, x_t^B\right) - c(q_t,S)\right] + \left[1-\mu(n_t)\right]u\left(0, x_t^B\right) + \Lambda u\left(0, x_t^S\right) - AH_t - n_t\kappa\end{aligned}$$

where in the second line we used the definition of H in (78).

Before we turn to the derivation of the PVIC, it is useful to derive the expressions for the interest rate and real money balances. Combining (34) and (35) we get and expression for the nominal interest rate

$$R_t = \mu(n_t)\left[\frac{u_q(q_t, x_t^B)}{g_q\left(q_t, x_t^B, S\right)} - 1\right] + 1 \tag{79}$$

Also, combining (35), (36) and (79), we get

$$\mathcal{M}_t = \frac{R_t g\left(q_t, x_t^B, S\right)}{u_x(0, x_t^S)} = \frac{R_t g\left(q_t, x_t^B, S\right) S(1-\tau)}{A} \tag{80}$$

which defines the money demand equation.

In order to construct the PVIC we take the budget constraint of each of the four types of agents and sum over time, discounting by the multiplier of the budget constraint for the period. We then aggregate these expressions using the measures we obtained above. Since the multipliers for all four types are identical, this amounts to starting with the aggregate budget constraint of households which we derived in Appendix B.2, multiplying with the multiplier $A/p_t S(1-\tau_t)$ in every period and summing over time. Doing this yields

$$\begin{aligned}\sum_{t=0}^{\infty}\beta^t\Bigg[&\frac{A}{S(1-\tau_t)}X_t - AH_t + \frac{A}{S(1-\tau_t)}\mathcal{B}_t - \frac{A}{S(1-\tau_t)}\frac{R_{t-1}\mathcal{B}_{t-1}}{\pi_t} \\ &+\frac{A}{S(1-\tau_t)}\mathcal{M}_t - \frac{A}{S(1-\tau_t)}\frac{\mathcal{M}_{t-1}}{\pi_t}\Bigg] = 0\end{aligned}$$

where we defined $X_t \equiv x_t^B + \Lambda x_t^S$.

Using (32), we can write

$$\sum_{t=0}^{\infty}\beta^t\left[u_x(0,x_t^S)X_t - AH_t\right] + \sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\mathcal{B}_t + \sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\mathcal{M}_t \tag{81}$$
$$= \sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\frac{\mathcal{M}_{t-1}}{\pi_t} + \sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\frac{R_{t-1}\mathcal{B}_{t-1}}{\pi_t}$$

Substitute into the second summation on the left-hand-side in (81) the equilibrium condition (34) to yield

$$\sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\mathcal{B}_t = \sum_{t=0}^{\infty}\beta^{t+1}R_t u_x(0,x_{t+1}^S)\frac{\mathcal{B}_t}{\pi_{t+1}}$$

and this will cancel with the second term on the right-hand-side to yield

$$\sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\frac{R_{t-1}\mathcal{B}_{t-1}}{\pi_t} - \sum_{t=0}^{\infty}\beta^{t+1}u_x(0,x_{t+1}^S)\frac{R_t\mathcal{B}_t}{\pi_{t+1}}$$
$$= \quad u_x(0,x_0^S)\frac{R_{-1}\mathcal{B}_{-1}}{\pi_0}$$

on the right hand side.

Next, substitute into the third summation on the left-hand-side in (81) the equilibrium condition (34) to yield

$$\sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\mathcal{M}_t = \sum_{t=0}^{\infty}\beta^{t+1}\frac{R_t u_x(0,x_{t+1}^S)\mathcal{M}_t}{\pi_{t+1}}$$

and using (79) we get

$$\sum_{t=0}^{\infty}\beta^{t+1}\frac{u_x(0,x_{t+1}^S)\mathcal{M}_t}{\pi_{t+1}}\left\{\mu\left(n_t\right)\left[\frac{u_q(q_t,x_t^B)}{g_q\left(q_t,x_t^B,S\right)}-1\right]+1\right\} \tag{82}$$
$$= \quad \sum_{t=0}^{\infty}\beta^{t+1}\frac{u_x(0,x_{t+1}^S)\mathcal{M}_t}{\pi_{t+1}}\mu\left(n_t\right)\left[\frac{u_q(q_t,x_t^B)}{g_q\left(q_t,x_t^B,S\right)}-1\right] + \sum_{t=0}^{\infty}\beta^{t+1}\frac{u_x(0,x_{t+1}^S)\mathcal{M}_t}{\pi_{t+1}}$$

Now, the second term in (82) cancels with the first term on the right-hand-side of (81) to

yield

$$\begin{aligned}&\sum_{t=0}^{\infty}\beta^t u_x(0,x_t^S)\frac{\mathcal{M}_{t-1}}{\pi_t}-\sum_{t=0}^{\infty}\beta^{t+1}\frac{u_x(0,x_{t+1}^S)\mathcal{M}_t}{\pi_{t+1}}\\ =\;& u_x(0,x_0^S)\frac{\mathcal{M}_{-1}}{\pi_0}\end{aligned}$$

on the right-hand-side. Using (32) and (36) on the first term in (82) we get

$$\begin{aligned}&\sum_{t=0}^{\infty}\beta^t\frac{\beta u_x(0,x_{t+1}^S)\mathcal{M}_t}{\pi_{t+1}}\mu\left(n_t\right)\left[\frac{u_q(q_t,x_t^B)}{g_q\left(q_t,x_t^B,S\right)}-1\right]\\ =\;&\sum_{t=0}^{\infty}\beta^t g\left(q_t,x_t^B,S\right)\mu\left(n_t\right)\left[\frac{u_q(q_t,x_t^B)}{g_q\left(q_t,x_t^B,S\right)}-1\right]\end{aligned}$$

on the left-hand-side.

To summarize we simplified (81) to

$$\begin{aligned}&\sum_{t=0}^{\infty}\beta^t\left\{u_x(0,x_t^S)X_t-AH_t+\mu\left(n_t\right)g\left(q_t,x_t^B,S\right)\left[\frac{u_q(q_t,x_t^B)}{g_q\left(q_t,x_t^B,S\right)}-1\right]\right\}\\ &=u_x(0,x_0^S)\left[\frac{R_{-1}\mathcal{B}_{-1}+\mathcal{M}_{-1}}{\pi_0}\right]\end{aligned}\tag{83}$$

using equilibrium conditions (32), (34), (35) and (36). This leaves equilibrium conditions (33), (37), (38) and (39). Out of these, (38) is redundant since we used the households' budget constraint to derive the PVIC. The rest of the remaining equations will have to be additional constraints on the Ramsey planner.

Given the allocations that are found from this problem, R_t will follow from (79) and the tax rate will follow from inverting (32).

B.4 Proof of Proposition 4

To solve the problem defined in Proposition 3, we do a few modifications. First, as Aruoba and Chugh (2010) shows for a similar problem, the Ramsey planner in this model does not have any intertemporal margins to manipulate. In any case we are just interested in the long-run outcome. and we drop all time subscripts. To ease notation, we use the shorthands defined in (48) and drop all arguments of remaining functions. We also realize that the FOC

of this problem with respect to H_t simply yields

$$-A - \xi A + S\nu = 0 \tag{84}$$

which can be solved for

$$\nu = \frac{(1+\xi)\,A}{S} \tag{85}$$

and we use this directly in all the equations below.

(49) follows from the FOC with respect to x^B

$$\begin{aligned}\mu(n)\,u_x^M + [1-\mu(n)]\,u_x^U + \xi\left\{u_x^S + \mu(n)\left[g_x\left(\frac{u_q^M}{g_q}-1\right) + \frac{g}{g_q^2}\left(u_{qx}^M g_q - u_q^M g_{qx}\right)\right]\right\} \\ -\frac{(1+\xi)\,A}{S} + \lambda\left\{\mu(n)\,u_{xx}^M + [1-\mu(n)]\,u_{xx}^U\right\} + \eta\mu(n)\,g_x + \iota\mu(n)\left(u_{qx}^M - g_{qx}\right) = 0\end{aligned} \tag{86}$$

(50) follows from the FOC with respect to x^S

$$\Lambda u_x^S + \xi\Lambda\left(u_{xx}^S x^S + u_x^S\right) - \Lambda\frac{(1+\xi)\,A}{S} - \lambda u_{xx}^S = 0 \tag{87}$$

(51) follows from the FOC with respect to q

$$\begin{aligned}\mu(n)\left(u_q^M - c_q\right) + \xi\mu(n)\left[g_q\left(\frac{u_q^M}{g_q}-1\right) + \frac{g}{g_q^2}\left(u_{qq}^M g_q - u_q^M g_{qq}\right)\right] + \lambda\mu(n)\,u_{xq}^M \\ +\eta\mu(n)\left(g_q - c_q\right) + \iota\mu(n)\left(u_{qq}^M - g_{qq}\right) = 0\end{aligned} \tag{88}$$

and (52) follows from the FOC with respect to n

$$\begin{aligned}\mu'(n)\left[u^M - c\right] - \kappa - \mu'(n)\,u^U + \xi\mu'(n)\,g\left(\frac{u_q^M}{g_q}-1\right) + \lambda\mu'(n)\left(u_x^M - u_x^U\right) \\ +\eta\left[\mu'(n)\,(g-c) - \kappa\right] + \iota\mu'(n)\left(u_q^M - g_q\right) = 0\end{aligned} \tag{89}$$

(53) is the PVIC with steady state imposed and (54) is the resource constraint. Finally, (55) and (56) are the UT and FE conditions, which are equality constraints and (57) is the complementary slackness condition that arises from the ZLB.

Figure 1: Five Key Facts

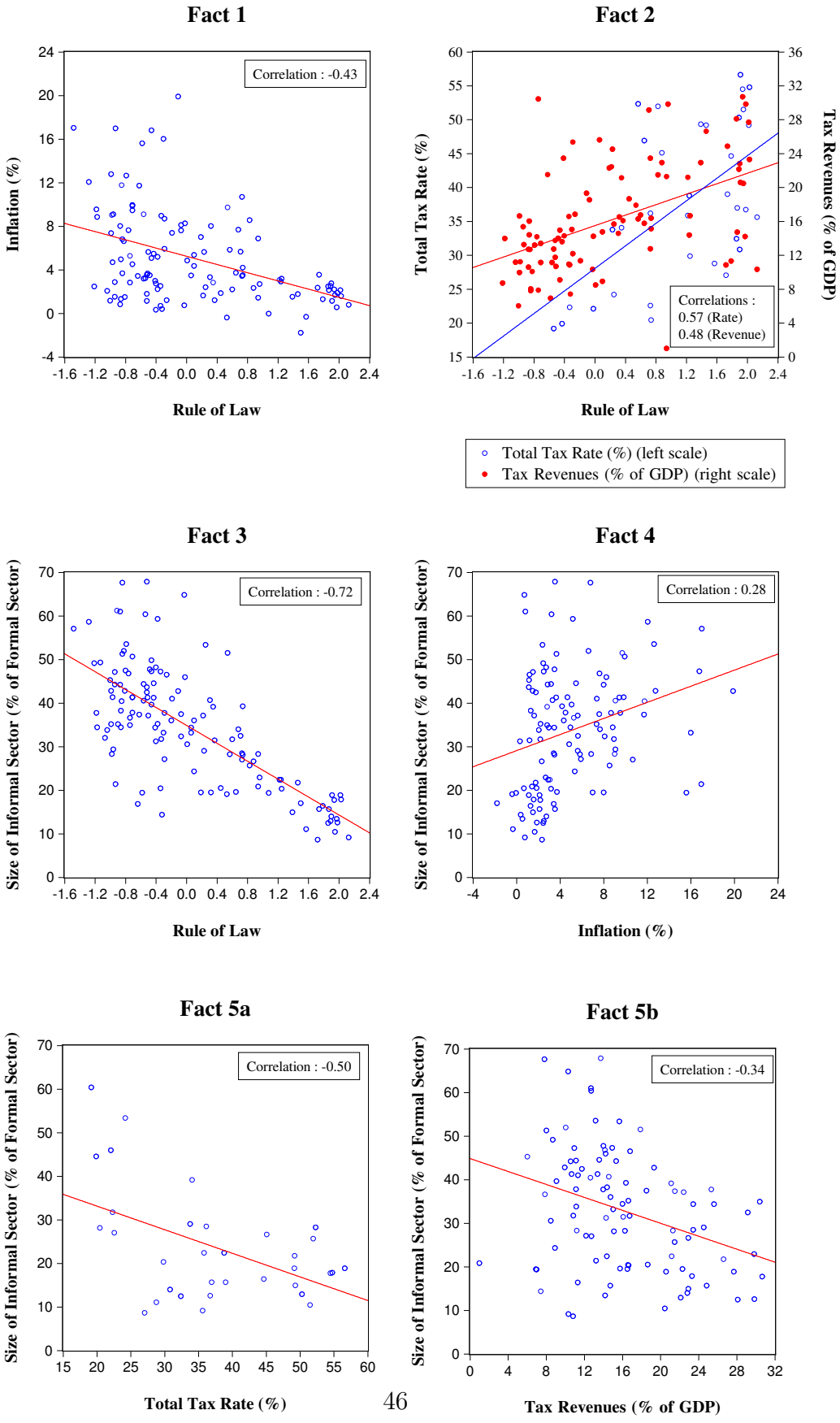
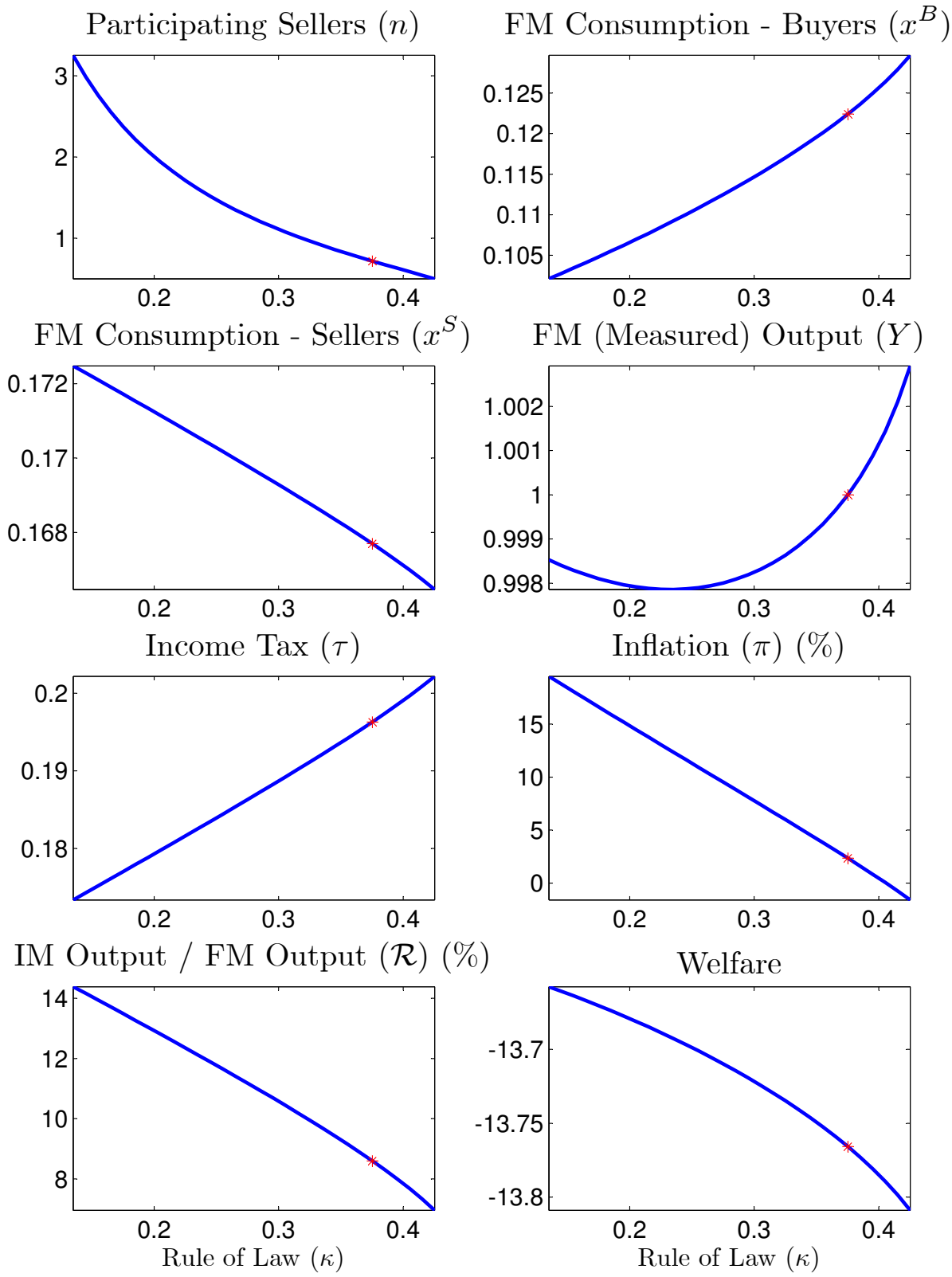
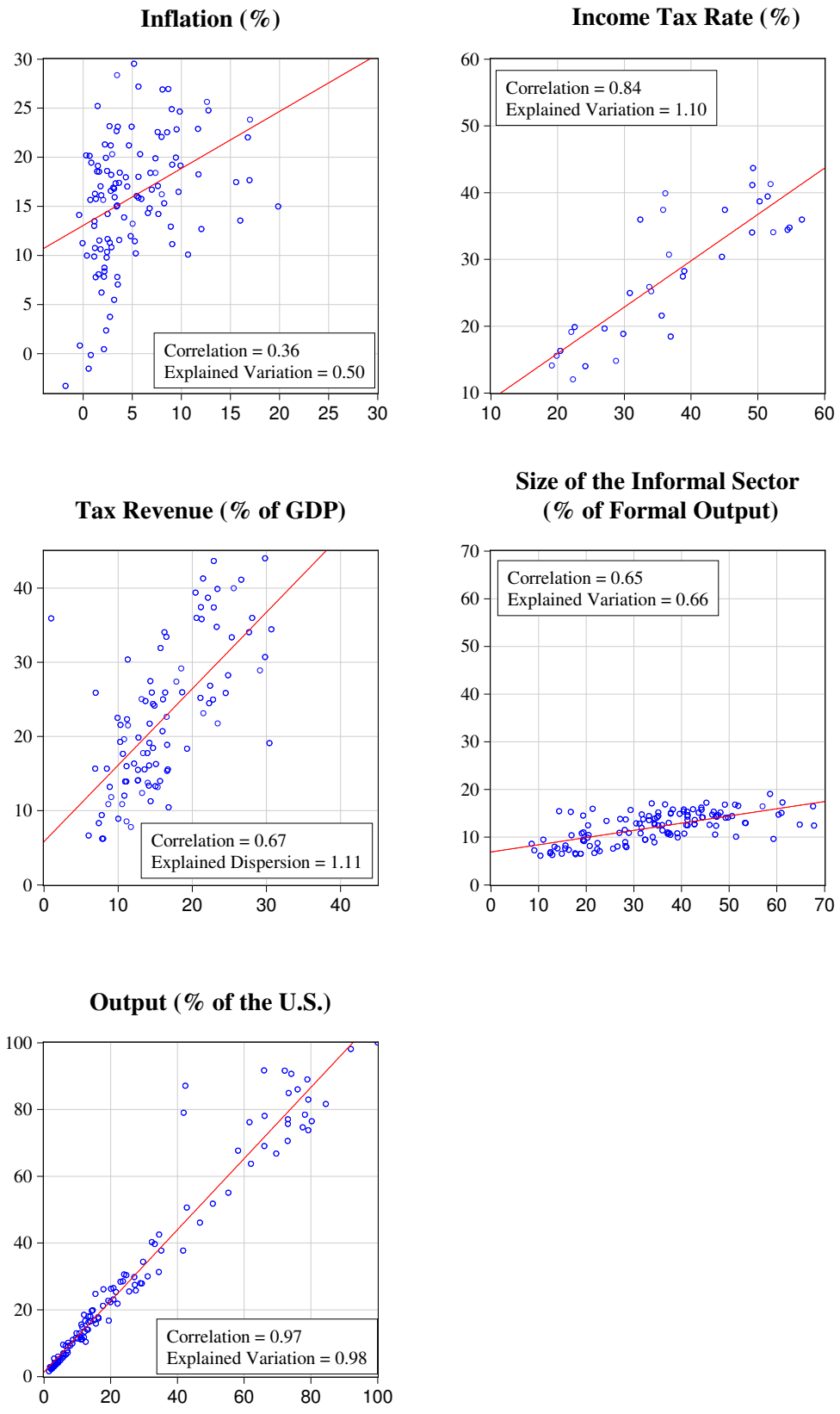


Figure 2: Numerical Comparative Statics



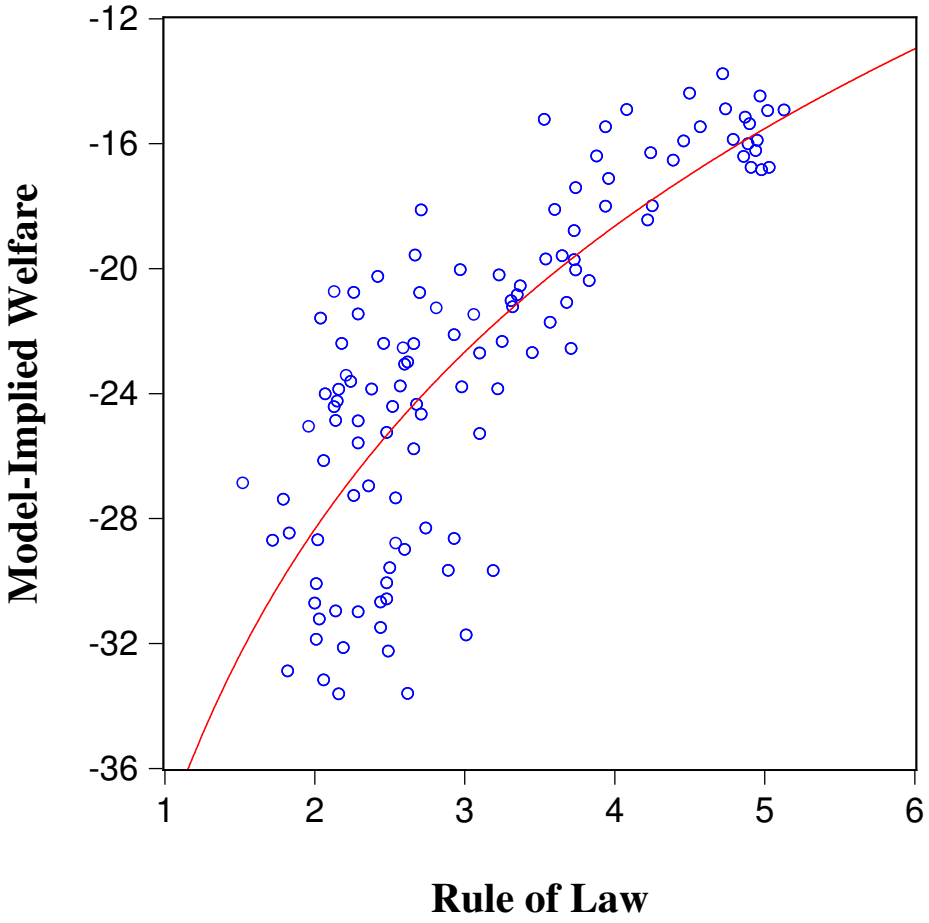
Notes: The red stars in each panel correspond to the calibrated value for the United States.

Figure 3: Cross Country Exercise - Key Variables, Data vs. Model



Notes: In each panel the x-axis shows the data and the y-axis reports the results from the model. Each figure is a square so that values above (below) the 45-degree line show the model predicts higher (lower) values relative to the data.

Figure 4: Welfare Across Countries



Notes: Figure shows welfare for each country in the cross-country exercise versus their rule of law. Welfare is calculated as defined in (40) and it is in arbitrary units. The x-axis shows Rule of Law plus three.

Table 1: Simple Correlations between Institutions, Government Policy and the Size of the Informal Sector

(a) Facts 1, 2 and 3

	Correlations of ...			
	Rule of Law (GM)	Irregular Payments (WEF)	Property Rights (WEF)	Freedom from Corruption (HF)
... with Rule of Law	-	0.87	0.86	0.94
Fact 1				
... with Inflation	**-0.43**	-0.57	-0.57	-0.41
Fact 2				
... with Total Tax Rate	**0.57**	0.47	0.44	0.50
... with Tax Revenue	**0.48**	0.61	0.40	0.48
Fact 3				
... with Informal Size (Schneider)	**-0.72**	-0.58	-0.65	-0.65
... with Informal Size (WEF)	-0.83	-0.72	-0.71	-0.75
... with Unrep. Wages Profits	-0.75	-0.77	-0.70	-0.77
... with Tax Evasion	-0.82	-0.77	-0.76	-0.81

(b) Facts 4 and 5

	Correlations of ...			
	Size (Schneider)	Size (WEF)	Unreported Wages / Profits	Tax Evasion
... with Informal Size (Schneider)	-	0.78	0.66	0.63
Fact 4				
... with Inflation	**0.28**	0.59	0.54	0.60
Fact 5				
... with Total Tax Rate	**-0.50**	-0.51	-0.41	-0.20
... with Tax Revenue	**-0.34**	-0.50	-0.51	-0.35

Note: The numbers in boldface correspond to the "headline" correlations used in Figure 1.

Table 2: Determinants of Government Policies

Dependent Variable : Inflation			
	(1)	(2)	(3)
Rule of Law	-	-1.82 (*)	-
Rule of Law (instrumented)	-	-	-1.43 (*)
Productivity	-0.07 (*)	-0.01	-
Government Exp.	0.04	0.02	-
R^2	0.17	0.20	0.07
N	118	118	104
Dependent Variable : Tax Rate			
	(4)	(5)	(6)
Rule of Law	-	6.15 (*)	-
Rule of Law (instrumented)	-	-	9.79 (*)
Productivity	0.09	-0.07	-
Government Exp.	0.90 (*)	0.88 (*)	-
R^2	0.71	0.75	0.43
N	34	34	34
Dependent Variable : Tax Revenue			
	(7)	(8)	(9)
Rule of Law	-	2.93 (*)	-
Rule of Law (instrumented)	-	-	4.06 (*)
Productivity	-0.00	-0.08	-
Government Exp.	0.45 (*)	0.41 (*)	-
R^2	0.44	0.48	0.26
N	97	97	84

Note: (*) indicates significance at 5% level. All regressions include a constant whose value is not reported. The instrumented rule of law uses Hall and Jones (1999) instruments and the first stage has an R^2 of 0.60. See text for details.

Table 3: Determinants of Size of Informal Sector

	(1)	(2)	(3)	(4)	(5)	(6)
Inflation	0.92 (*)	-0.11	-	-	-	-
Tax Rate	-	-	-0.54 (*)	-0.00	-	-
Tax Revenue	-	-	-	-	-0.74 (*)	0.01
Rule of Law	-	-10.43 (*)	-	-13.80 (*)	-	-10.18 (*)
R^2	0.08	0.52	0.25	0.77	0.12	0.52
N	118	118	34	34	97	97

	(7)	(8)	(9)
Rule of Law	-	-6.82 (*)	-
Rule of Law (instrumented)	-	-	-11.29 (*)
Productivity	-0.33 (*)	-0.13	-
Government Exp.	-0.05	0.02	-
R^2	0.49	0.53	0.42
N	118	118	104

Note: All regressions have size of the informal market as the dependent variable. (*) indicates significance at 5% level. All regressions include a constant whose value is not reported. The instrumented rule of law uses Hall and Jones (1999) instruments and the first stage has an R^2 of 0.60. See text for details.

Table 4: Main Results

	Data	**Benchmark**	**Constant κ**	**Only κ**	**Constant S**	**Only G**
Correlation of Model-Based Measures and Data						
Inflation	-	0.36	-0.11	0.43	0.37	-0.09
Tax Rate	-	0.84	0.82	0.56	0.84	0.82
Tax Revenue	-	0.67	0.65	0.48	0.67	0.65
$\mathcal{R}$	-	0.64	0.23	0.72	0.65	0.33
Output	-	0.97	0.97	0.61	0.45	0.42
Fraction of Cross-Country Dispersion Explained by Model						
Inflation	-	0.50	1.41	0.60	0.50	1.24
Tax Rate	-	1.10	1.04	0.13	1.10	1.04
Tax Revenue	-	1.11	0.96	0.13	1.12	0.97
$\mathcal{R}$	-	0.66	0.26	0.46	0.67	0.28
Output	-	0.98	0.97	0.00	0.01	0.01
Facts - Data and Model-Implied						
1 - $corr(\kappa, \pi)$	-0.43	-0.60	0.51	-1.00	-0.62	0.47
2a - $corr(\kappa, \tau)$	0.57	0.46	0.39	1.00	0.46	0.39
2b - $corr(\kappa, \tau)$	0.48	0.66	0.60	1.00	0.67	0.60
3 - $corr(\kappa, \mathcal{R})$	-0.72	-0.89	-0.35	-1.00	-0.90	-0.47
4 - $corr(\pi, \mathcal{R})$	0.28	0.19	-0.95	1.00	0.22	-1.00
5a - $corr(\tau, \mathcal{R})$	-0.50	-0.80	-1.00	-1.00	-0.80	-1.00
5b - $corr(\tau, \mathcal{R})$	-0.34	-0.88	-0.97	-1.00	-0.88	-1.00

Notes: The table reports various statistics from different versions of the model where one or more of the exogenous variables κ, S and G are allowed to vary across countries. $\mathcal{R}$ denotes the ratio of informal activity to formal activity, π denotes inflation and τ denotes taxes. For Facts 2 and 5, (a) refers to the statistics computed using tax rates and (b) refers to the statistics computed using tax revenues. Throughout the table, results with tax rates uses 34 countries, results with tax revenues uses 97 countries and all others use all 118 countries, with the exception of the version with constant κ for which no solution could be obtained for 2 countries and the version with only G for which no solution can be obtained for 3 countries.

Table 5: Results for Subsamples of Countries

	Benchmark	**Low Institutions**	**Low Output**	**Other Countries**
	(118)	(19)	(66)	(49)
Correlation of Model-Based Measures and Data				
Inflation	0.36	0.16	0.20	0.46
Tax Rate	0.84	-	0.09	0.79
Tax Revenue	0.67	0.36	0.66	0.58
$\mathcal{R}$	0.64	0.07	0.18	0.63
Output	0.97	0.98	0.96	0.92
Fraction of Cross-Country Dispersion Explained by Model				
Inflation	0.50	0.34	0.38	0.68
Tax Rate	1.10	-	1.44	1.13
Tax Revenue	1.11	1.43	0.99	0.83
$\mathcal{R}$	0.66	0.46	0.52	0.50
Output	0.98	1.00	1.06	0.98
Facts - Data and Model-Implied				
	Data / Model	Data / Model	Data / Model	Data / Model
1 - $corr(\kappa, \pi)$	-0.43 / -0.60	-0.36 / -0.26	-0.18 / -0.35	-0.54 / -0.67
2a - $corr(\kappa, \tau)$	0.57 / 0.46	- / -	1.00 / 0.16	0.36 / 0.23
2b - $corr(\kappa, \tau)$	0.48 / 0.66	0.33 / 0.06	0.43 / 0.44	0.32 / 0.26
3 - $corr(\kappa, \mathcal{R})$	-0.72 / -0.89	-0.27 / -0.15	-0.27 / -0.64	-0.77 / -0.82
4 - $corr(\pi, \mathcal{R})$	0.28 / 0.19	-0.30 / -0.91	-0.04 / -0.50	0.43 / 0.13
5a - $corr(\tau, \mathcal{R})$	-0.50 / -0.80	- / -	-0.16 / -0.60	-0.17 / -0.74
5b - $corr(\tau, \mathcal{R})$	-0.34 / -0.88	-0.29 / -0.98	-0.16 / -0.90	-0.20 / -0.72

Notes: The table reports various statistics from different subsets of countries. The number in parentheses in the top row report the number of countries in the particular sample. 'Low institutions' refers to countries with Rule of Law in the lowest quintile and 'low output' refers to countries that have less than 20% of the output of the U.S. 'Other countries' are those that satisfy neither of the criteria. '-' indicate subsamples where no tax rate data was available for the countries. See also the notes to Table 4.

Table 6: Results for Subsamples of Countries

	Benchmark	**Free**	**Not Free**
	(118)	(58)	(27)
Correlation of Model-Based Measures and Data			
Inflation	0.36	0.32	0.24
Tax Rate	0.84	0.83	-
Tax Revenue	0.67	0.69	0.37
$\mathcal{R}$	0.64	0.75	0.47
Output	0.97	0.99	0.95
Fraction of Cross-Country Dispersion Explained by Model			
Inflation	0.50	0.45	0.16
Tax Rate	1.10	1.12	-
Tax Revenue	1.11	1.08	0.79
$\mathcal{R}$	0.66	0.58	0.45
Output	0.98	0.98	1.32
Facts - Data and Model-Implied			
	Data / Model	Data / Model	Data / Model
1 - $corr(\kappa, \pi)$	-0.43 / -0.60	-0.43 / -0.60	-0.37 / -0.57
2a - $corr(\kappa, \tau)$	0.57 / 0.46	0.54 / 0.43	- / -
2b - $corr(\kappa, \tau)$	0.48 / 0.66	0.38 / 0.49	-0.01 / 0.65
3 - $corr(\kappa, \mathcal{R})$	-0.72 / -0.89	-0.81 / -0.88	-0.60 / -0.78
4 - $corr(\pi, \mathcal{R})$	0.28 / 0.19	0.21 / 0.15	0.43 / -0.07
5a - $corr(\tau, \mathcal{R})$	-0.50 / -0.80	-0.46 / -0.79	- / -
5b - $corr(\tau, \mathcal{R})$	-0.34 / -0.88	-0.36 / -0.83	0.21 / -0.91

Notes: The table reports various statistics from different subsets of countries. The number in parentheses in the top row report the number of countries in the particular sample. 'Free' refers to countries with Freedom House ratings between 1.0 and 2.5 (inclusive) and 'Not Free ' refers to countries with ratings 5.5 and above. '-' indicate subsamples where no tax rate data was available for the countries. See also the notes to Table 4.

Table 7: Countries and Key Variables

	Country	Rule of Law	Inflation	Tax Rate	Tax Revenue	Size
1	Albania	-0.86	4.96	-	16.02	34.43
2	Algeria	-0.74	2.82	-	30.43	34.90
3	Argentina	-0.29	5.93	-	12.18	27.13
4	Armenia	-0.48	3.49	-	14.00	47.73
5	Australia	1.90	2.76	30.83	22.82	13.97
6	Austria	1.95	1.67	51.50	20.46	10.43
7	Azerbaijan	-0.87	0.83	-	12.71	61.00
8	Bangladesh	-0.74	5.27	-	7.87	36.60
9	Belgium	1.46	1.77	49.20	26.63	21.73
10	Benin	-0.40	2.73	-	-	48.20
11	Bhutan	0.01	4.85	-	8.49	30.53
12	Bolivia	-0.52	3.53	-	13.72	67.83
13	Botswana	0.68	7.68	-	-	33.97
14	Brazil	-0.19	7.39	-	11.34	41.00
15	Bulgaria	-0.07	7.61	-	18.53	37.43
16	Burkina Faso	-0.52	1.78	-	11.74	42.43
17	Burundi	-1.18	9.53	-	13.97	37.73
18	Cambodia	-0.84	3.68	-	8.02	51.27
19	Cameroon	-1.04	2.05	-	11.17	33.80
20	Canada	1.87	2.12	36.96	14.72	15.67
21	Central African Rep	-1.00	1.16	-	6.03	45.27
22	Chad	-0.94	1.51	-	-	47.10
23	Chile	1.25	3.17	29.85	16.65	20.33
24	China	-0.32	0.39	-	7.42	14.37
25	Hong Kong	1.50	-1.79	-	-	16.97
26	Colombia	-0.71	9.45	-	13.40	41.27
27	Congo	-1.21	2.45	-	8.70	49.13
28	Costa Rica	0.73	10.69	22.57	12.76	27.00
29	Cte D'Ivoire	-0.94	2.87	-	15.34	44.23
30	Croatia	0.06	3.47	-	25.62	34.33

Table 7: Countries and Key Variables

	Country	Rule of Law	Inflation	Tax Rate	Tax Revenue	Size
31	Czech Republic	0.65	3.73	-	15.76	19.60
32	Denmark	1.94	2.18	54.49	30.69	17.73
33	Dominican Rep	-0.30	16.01	-	15.05	33.20
34	Egypt	0.10	4.35	-	14.74	36.00
35	El Salvador	-0.34	2.50	-	10.95	47.23
36	Estonia	0.74	4.18	-	16.37	39.23
37	Ethiopia	-0.51	3.66	-	10.67	41.27
38	Fiji	-0.41	2.97	-	23.44	34.33
39	Finland	2.03	1.59	54.77	23.31	17.83
40	France	1.39	1.52	49.32	22.92	14.90
41	Georgia	-0.84	6.79	-	7.82	67.63
42	Germany	1.79	1.29	44.65	11.32	16.37
43	Ghana	-0.11	19.89	-	19.32	42.73
44	Greece	0.73	3.42	36.17	23.43	28.47
45	Guatemala	-0.82	6.62	-	10.06	51.93
46	Haiti	-1.48	17.02	-	-	57.03
47	Honduras	-0.71	9.94	-	-	50.67
48	Hungary	0.83	8.56	51.95	21.48	25.67
49	India	0.10	5.37	-	8.91	24.30
50	Indonesia	-0.93	16.97	-	13.24	21.37
51	Iran	-0.58	15.61	-	6.92	19.40
52	Ireland	1.74	3.55	39.01	24.85	15.63
53	Israel	0.96	2.69	-	29.84	22.87
54	Italy	0.88	2.33	45.09	22.92	26.60
55	Jamaica	-0.29	8.68	-	25.34	37.70
56	Japan	1.57	-0.33	28.77	-	11.03
57	Jordan	0.45	1.85	-	18.66	20.50
58	Kazakhstan	-0.87	8.02	-	10.58	44.17
59	Kenya	-0.99	7.37	-	16.64	35.13
60	Korea	0.74	3.50	20.43	15.13	28.13

Table 7: Countries and Key Variables

	Country	Rule of Law	Inflation	Tax Rate	Tax Revenue	Size
61	Kuwait	0.94	1.43	-	1.00	20.80
62	Kyrgyzstan	-0.85	11.76	-	12.66	40.43
63	Latvia	0.31	3.32	-	14.57	40.63
64	Lesotho	-0.07	8.10	-	-	32.33
65	Lithuania	0.37	1.21	-	16.09	31.43
66	Macedonia	-0.38	2.22	-	-	35.17
67	Madagascar	-0.56	9.09	-	-	40.53
68	Malaysia	0.60	2.18	-	16.75	31.63
69	Maldives	-0.40	0.33	-	14.28	31.23
70	Mali	-0.52	1.13	-	-	43.63
71	Mauritania	-0.50	5.64	-	-	37.10
72	Mexico	-0.33	8.95	22.31	10.86	31.70
73	Moldova	-0.46	16.79	-	14.94	47.27
74	Mongolia	0.19	7.01	-	22.30	19.47
75	Morocco	0.22	1.63	-	22.42	37.13
76	Mozambique	-0.71	9.85	-	-	41.33
77	Namibia	0.71	5.65	-	29.15	32.47
78	Nepal	-0.46	5.06	-	9.08	39.63
79	Netherlands	1.89	2.50	50.29	22.13	12.90
80	New Zealand	1.98	1.87	36.73	29.84	12.57
81	Nicaragua	-0.76	7.62	-	14.17	46.77
82	Niger	-0.81	1.49	-	-	42.77
83	Nigeria	-1.28	12.05	-	-	58.63
84	Norway	2.02	2.13	49.17	27.69	18.83
85	Oman	1.08	-0.04	-	7.01	19.37
86	Pakistan	-0.71	4.50	-	11.13	37.80
87	Panama	-0.03	0.73	-	10.32	64.83
88	Papua N Guinea	-0.62	11.72	-	21.51	37.33
89	Paraguay	-0.96	9.09	-	-	29.33
90	Peru	-0.54	3.22	19.16	12.74	60.37

Table 7: Countries and Key Variables

	Country	Rule of Law	Inflation	Tax Rate	Tax Revenue	Size
91	Philippines	-0.43	5.48	19.89	13.58	44.50
92	Poland	0.57	5.83	52.34	16.27	28.23
93	Portugal	1.22	3.06	35.85	21.18	22.37
94	Rwanda	-0.97	4.68	-	-	41.30
95	Saudi Arabia	0.53	-0.40	-	-	19.07
96	Senegal	-0.26	1.21	-	16.82	46.47
97	Sierra Leone	-0.99	12.77	-	9.95	42.80
98	Singapore	1.97	0.56	-	14.20	13.40
99	Slovakia	0.32	8.00	-	16.52	19.47
100	Slovenia	0.94	6.86	-	21.29	28.27
101	Solomon Islands	-1.17	8.84	-	-	34.40
102	South Africa	0.23	5.64	33.77	24.53	29.00
103	Spain	1.24	2.90	38.79	14.37	22.40
104	Sri Lanka	-0.02	8.26	22.08	14.24	45.90
105	Sweden	1.91	1.14	56.61	20.59	18.87
106	Switzerland	2.13	0.78	35.61	10.33	9.13
107	Syria	-0.34	0.68	-	16.57	20.43
108	Tanzania	-0.38	5.19	-	-	59.30
109	Thailand	0.25	2.40	24.18	15.69	53.37
110	Togo	-0.86	1.32	-	14.39	38.23
111	Tunisia	0.35	2.83	34.06	21.13	39.13
112	Uganda	-0.56	3.18	-	11.14	44.37
113	Ukraine	-0.79	12.63	-	13.20	53.50
114	U.K.	1.86	2.46	32.42	28.11	12.47
115	U.S.A.	1.72	2.35	27.05	10.84	8.60
116	Uruguay	0.54	9.73	-	17.89	51.47
117	Vietnam	-0.64	3.42	-	-	16.80
118	Yemen	-0.98	9.03	-	11.22	28.30
	Average	**0.10**	**5.06**	**37.45**	**16.14**	**33.76**
	Minimum	**-1.48**	**-1.79**	**19.16**	**1.00**	**8.60**
	Maximum	**2.13**	**19.89**	**56.61**	**30.69**	**67.83**