Research Reports
Kansantaloustieteen laitoksen tutkimuksia, No. 119:2009
Dissertationes Oeconomicae

JENNI PÄÄKKÖNEN
Essays on Illegal Activities, Institutions and Economic Growth

ISBN 978-952-10-5346-7 (nid)
ISSN 0357-3257
Acknowledgements

This dissertation was partly written while I worked as a research assistant in the Department of Economics at the University of Helsinki. I am very grateful to Professors Essi Eerola, Vesa Kanniainen and Ulla Lehmijoki for their support and guidance in writing the thesis. I wish to extend my thanks to Professor Pertti Haaparanta and Academy Professor Erkki Koskela for their advice and constructive suggestions. As well, I wish to thank Docent Kari Alho and Professor Mika Linden, the pre-examiners of my dissertation, for their helpful suggestions and efficient work. I have also greatly benefitted from discussions with and support from my other colleagues at the University of Helsinki; for purposes of brevity I cannot name you all, but I include everyone who worked in the Department of Economics during 2003-2008. I have also greatly benefitted from the numerous comments of various seminar instructors and participants in the Hecer and FDPE. I also like to thank Director Otto Toivanen for his insightful advice concerning science and the general prospects for my career.

I have been working at the Bank of Finland’s Institute for Economies in Transition since autumn 2008. I am very grateful to its former head, Pekka Sutela, for giving me this opportunity. Together with our current head, Iikka Korhonen, he advised me and most of all gave me a chance to finish this dissertation on sovereignty. I wish to extend my thanks to all my colleagues here at BOFIT, who made me feel instantly welcome after my arrival and offered constructive comments and encouragement in the later stages of the work. In particular, I thank my colleagues at the China desk, who were willing to bear more than their fair share of our workload in order to let me perfect my dissertation without extra stress from other tasks.

Financial support from the Yrjö Jahnsson Foundation and Suomen Konkordialliitto is gratefully acknowledged.

Finally, I wish to thank my spouse Jyrki for his continuous support, encouragement, waterproof reasoning, and sound advice during the times when the research seemed to stall, and, in particular, during the times of chaos. I also owe much gratitude to my family and friends for being there whenever I needed them. I dedicate this thesis to all who are close to me, and to those relatives who did not live to see the day of my graduation.
Introduction

November 12, 2009

Contents

1 General Background 1

2 Background for the Essays 4
  2.1 Optimal Law Enforcement in the Presence of Organized Crime . . . 5
  2.2 Do The Catholic and Protestant Countries Differ by Their Tax Morale? 6
  2.3 Economic Freedom as a Driver for Growth in Transition . . . . . . 8
  2.4 Economic Growth and Demographic Transition: A Club Approach . 11

3 Summary of the Essays 13
  3.1 Optimal Law Enforcement in the Presence of Organized Crime . . . 13
  3.2 Do The Catholic and Protestant Countries Differ by Their Tax Morale? 14
  3.3 Economic Freedom as a Driver for Growth in Transition . . . . . . 15
  3.4 Economic Growth and Demographic Transition: A Club Approach . 16
1 General Background

One of the areas in economics about which the economists seem to disagree most is the field of public economics, and in particular, the role and the functions of a government. Two extremist views compete: one emphasizes government working for the public interest to provide value for the citizens, while another regards government mainly as a workhorse for private interests. The proponents of the latter view favor the ideal that the government should interfere with the markets as little as possible, whereas the proponents of the former view note that there are areas where efficient outcomes are not attainable or where market outcomes are not socially desirable. In this case, the government’s role is to alleviate the market failures, while minimizing the distortions its interventions create in the markets. Even if the government is expected to provide public goods, there is a somewhat troublesome trade-off between providing enough public goods and distorting the markets as little as possible.

As the sole legitimate authority, the government has the right to define the rules and laws as well as to enforce them. Above all, it has the right to use coercion and force. Clearly, such rights put the government in a dominant position with regard to the citizens, giving government officials the incentive and the opportunity to (mis-)use their authority for private interests. Thereby it is not necessarily the case that both the size of the government or its regulatory actions are greater in a regime where government is held as optimizing the public rather than private interest. With respect to regulation, two extremes arise: from too little regulation to too much of it. If the government does not function or ceases to exist, the state falls into anarchy or chaos (Somalia). If it regulates too much, it will completely suffocate private activities, which might be considered extralegal (the former Soviet Union).

The role and the functions of the government were already debated by Plato in his Republic approximately 2400 years ago. In the Republic Plato laid down his ideals of morality, justice, governance and the role of the state in a society. On the other hand, quite recently Maddison (1991) listed 11 key long-run forces of capitalist development running from institutions to physical capital. While many of Plato’s ideals serve as ideals even today, they seem to interact with the forces of capitalist
development. This thesis debates these two seemingly unrelated issues, showing that the morale, governance and economic well-being are closely connected. This in fact in the current economic crisis seems intuitive and topical.

The four essays in this thesis debate some of the same ideals that were debated by Plato; I scrutinize the role of social norms and tax morale in the coexistence of legal and illegal production; the need for law enforcement in the presence of organized crime; the measurement of the size of the shadow economy; the measurement of material well-being; the role of public policy and institutions in economic growth and material well-being; and demographic transition as an impetus for economic growth. The government’s interventionist policies and the question of how to measure production and promote economic well-being are present in all of these essays.

The first two essays are in the nexus of public economics and law and economics, utilizing the apparatus familiar from industrial economics and measurement theory, while the third and fourth essays build more on what is traditionally regarded as empirical macroeconomics. The first two essays evaluate the dynamics and causes of the shadow economy. The first essay presents a model where the government and its challenger, an extralegal authority, collect rents from legal and illegal entities. The optimal policies of the government and the best responses from the challenger are evaluated. The second essay tackles the complex task of quantifying non-observables using structural equation techniques.

The last two essays view the dynamics of economic well-being. Maddison’s (1991) Dynamic Forces in Capitalist Development lists the following long-run forces of capitalist development: 1) institutions, 2) demographic change, 3) role of government, 4) structural change, 5) human capital, 6) physical capital, 7) technical progress, 8) growth of labor supply, 9) natural resources, 10) international trade and 11) colonialism. The last two essays debate the first six of them and the first three in particular. The third essay considers the nexus of institutions, transition from socialism to capitalism, role of government in this transition and structural change. It evaluates their role, but also role of human capital and physical capital as engines of growth during transition. The third essay also reviews the issue of measurement but contrasts two common ways to measure economic well-being: real GDP per
capita and real GDP per worker. The fourth essay reviews the old question put forth by Malthus’s *Essay on the Principle of Population* in the 18th century, i.e., the question of how demographic change affects well-being, and whether it is plausible to expect that poor countries will achieve the same standard of living as the rich ones. In the last two essays, by using frontier techniques in econometrics, we seek an answer to the question of how to best promote the economic well-being of transition economies and underdeveloped nations. In both of these essays, we pay attention to the non-linearities and differences in models across a cross-section of countries.

The first essay assumes that the government’s policies might promote illegal activity. It takes a look at an economy where the presence of underground production gives an incentive for an extralegal authority to extract rents from vulnerable producers. The paper evaluates the interaction between the (Leviathan) government and the mafia, and pays attention to the law enforcement of underground production. The main questions this paper asks are the following: What are the optimal policies of the government, and how do these affect legal and illegal production? What is the best response by the mafia? Should a revenue-maximizing government enforce laws, and what are the implications for welfare? One of the main findings is that the producers all suffer from an enforcement policy.

The second paper regards the illegal activity as an entrepreneur’s response to regulatory setting. It tackles the issue of the measurement of the size of the shadow economy, which is problematic since there are no direct ways to measure hidden activities. To formulate policies it is essential to know the approximate magnitude of these activities and in particular what drives illegal economic activity; is it the tax burden, excess regulation, corruption or a weak legal environment? In this paper we propose an additional explanation for tax evasion and shadow production, namely cultural factors manifested by religion as determinants of tax morality. We evaluate whether Catholic and Protestant countries differ in their tax morale and whether this is seen in the size of the shadow economy. According to our findings, this might not be the case.

The third paper contributes to the literature discussing the role of the government in promoting economic growth. The common view holds that the quality of
government is critical for economic success. The evidence for this idea is, however, unequivocal in the context of economies in transition, creating a need to verify whether this view is true. As informal production makes up a big portion of the total production in these countries, we make an attempt to correct for this bias by using the real GDP per worker. By concentrating on those who are listed as working in the formal economy, we better measure the average productivity and the growth potential of these nations. Lastly, the paper scrutinizes some earlier, contradictory findings regarding the drivers of economic growth in transition. Our main result is that, given the complex relationship between economic growth and economic freedom, marketization has not necessarily been beneficial in terms of growth.

The last paper builds on traditional growth literature and revisits the debate on convergence clubs arising from demographic transition. We scrutinize the presence of demographic clubs and the ongoing process of demographic transition. We provide new evidence against the idea that countries within a club would converge over time. Instead, we propose that since the demographic transition is a dynamic process, one can expect countries to enter the last regime of stable, modern growth in stages. Prior to that, no convergence can be plausibly expected, and there might be countries that are left behind in a demographic trap, as proposed by Malthus (1798).

2 Background for the Essays

In this section, I discuss the background and literature of each essay. I briefly summarize my reading about the literature concentrating on those that are closest to my own contribution, which should help the reader to put my contribution in the context. For a broader and more in-depth debate on the literature, I encourage the reader to take a look at individual essays.
2.1 Optimal Law Enforcement in the Presence of Organized Crime

It appears that the economic crime is often organized, yet in realms of economics this phenomena was almost utterly neglected until mid 90’s and in particular until the frame-breaking contribution of Garoupa’s (2000). Since the complexity of the topic, there are controversies and even contradictory views on this phenomenon as will be manifested in what follows.

The *International Encyclopedia of the Social and Behavioral Science* defines organized crime as follows a) it has some manner of formalized structure whose primary objective is to obtain money through illegal activities; b) it has continuity over time; and c) it maintains its position through use of violence or threat of violence, corruption of public officials, and extortion. This definition is in stark contrast with the traditional, somewhat romanticized, view that for example Cosa Nostra in Italy emerged as a defense mechanism for impoverished rural peasants against their landlords (e.g. Allum and Sands, 2004).

While the strength of the government, or lack of it, has been highlighted by many authors\(^1\), the excessive regulation and the potential for illegal market are also important for the birth of organized crime (see Anderson 1995). Thereby the problem with organized crime is and has been more rampant in transition economies as they have undergone a major institutional reform from communism to market economies. According to Johnson and Kaufmann (2001), these economies experienced a growth of illicit activities and organized crime while the state was weak. It is well known, as soon as they have emerged, crime and corruption are difficult to root out.

Several papers have discussed the behavior of the criminal organization and optimal law enforcement. The seminal contribution is Garoupa’s (2000) with vertical structure between the principal (criminal organization) and agent (individuals), where the former extracts rents from the latter. Garoupa shows that there are cases where the presence of criminal organization reduces the need to enforce laws by the government, since the presence of crime group makes the illegal activity more

---

\(^1\)See for example Williams and Godson (2002), Skaperdas and Syropoulos (1995), and Frye and Shleifer (1997).
costly. This result has been replicated by many papers with different assumptions on the organization structure\(^2\), the information structure and with the possibility of the collusion between the police and the criminals\(^3\). What is interesting for sake of optimal law enforcement, they all contradict Becker’s (1968) famous result on maximal fine. Instead, they propose that under certain circumstances increasing policing and sanctions can generate higher crime rates. This view is sometimes repeated in public and it merits more discussion as to some extend it runs counter to intuition and common sense.

While the studies cited above narrowed the role of the mafia only to a harmful, rent-extracting organization, some studies have evaluated the possibility that the mafia would provide public goods. These studies regard the mafia as an authority competing with the government. For example Anderson and Bandiera (2000) present mafias as enforcement coalitions, which protect property from predators. Also Grossman (1995) and Alexeev et al. (2004) emphasize the role of the mafia as a provider of public goods. In the first paper, the private producer is better off as the government’s and the mafia’s competition increases the production of public goods, while in the latter the revenue-maximizing government may actually benefit from the presence of the mafia as the mafia makes the underground production costlier. In reality, the mafia’s public good is, however, merely defense from its own violence and not from an outside threat. Then the question is how does these results hold in more general setting where the mafia does not provide anything in exchange for the rent.

### 2.2 Do The Catholic and Protestant Countries Differ by Their Tax Morale?

On moral norms the seminal paper is due to Macaulay (1963), who suggests that people behave honestly because honesty is rewarded and defection punished in future transactions. Brekke et al. (2003) propose a model where a consumer’s self-image can only be improved by striving towards what she truly believes to be morally right. This idea is replicated by Johansson-Stenman and Martinsson.

\(^2\)For interesting paper with and endogenous gang structure see Mansour et al. (2006).
\(^3\)See Bowles and Garoupa (1997), Garoupa and Jellal (2007) and Kugler et al. (2005).
The theory of tax evasion goes back at least to Allingham and Sandmo (1972), but it was Baldry (1987) who suggested that there are moral costs to cheating and that moral feelings have to be incorporated in the theory of tax evasion. Following this line of argument, Gordon (1989) incorporates social norms into the evasion decision assuming that the utility cost to evasion increases in the proportion of taxpayers who do not evade. The experiment by Myles and Naylor (1996) confirm a tax evasion model where a social custom utility is derived when taxes are paid honestly and a conformity payoff from adhering to the standard pattern of social behavior. The experimental results of Orvaska and Hudson (2002) suggest that evasion is condoned by a large proportion of the population though people appear to be deterred from tax evasion by the consequences of being caught.

The interaction of tax ethics and tax rates has been evaluated by Reckers et al. (1994). One of their findings is that those who already have a low tax morale or do not condemn tax evasion seem to respond more in terms of noncompliance to the increases in tax rate. Reviewing literature Pilkington (2007) suggests that tax evasion is subject to group norm; if tax evasion is widespread one may cheat without feeling guilty. This view is verified by Feld et al. (2008) who evaluate the tax morale in Germany after German unification. They find that from 1990 to 1999 the initially higher East German tax morale converged to the lower West German levels. Lastly, other researches propose that there exist cultural differences regarding tax morale across countries. For example, Alm and Torgler (2006) suggest that northern Europeans have higher tax morale than southern Europeans.

There are numerous empirical studies on the economic effects of religion (cf. Iannaccone, 1998). Substantial differences exist between Catholicism and Protestantism in both the structure of their beliefs and their enforcement mechanisms. Arrunada (2004) argues that the catholic theology and practice facilitate personal transactions while protestantism favors values and types of moral and legal enforcement better adapted to impersonal trade. Guiso et al. (2003) find an empirical

\footnote{Falkinger (1995) notes that perceived fairness and equity of the political and economic system increases the bad conscience of evaders. Fortin et al. (2004) find that perceived unfair taxation may lead to increased tax evasion.}
relation between religion and people’s attitudes toward legal rules. In particular, they find that religious denomination affects attitudes toward legal norms, as cheating on taxes. On the causes of the shadow economy, Loyaza (1996) finds that the size of the shadow economy is positively related to the tax burden and negatively on the quality of government institutions. Johnson and Kaufmann (2001) find that the underground economy in transition economies has mainly been driven by excess regulation and corruption and weak legal environment and to a lesser extent by taxation.

While other methods have also been used, structural equation models have become a standard technique for estimating the size of the shadow economy. Its origins are in simultaneous equations models (Haavelmo 1943), while the treatment of latent variables was incorporated by Zellner (1970) and Jöreskog and Goldberger (1975). The existing empirical studies quantifying the size of the shadow economy have been evaluated critically by several authors. According to Hill (2002), there is typically no theory behind the estimated model. He also questions the identification of the latent variable and the reliability of the results, which are difficult to judge as there are no reliable alternative methods. Hills’ critique is echoed by Smith (2002), who also argues that as the size of the shadow economy must be calibrated by using "outside information", this "other source" for obtaining the benchmark becomes critical.

The most severe critique is due to Breusch (2005a,b) who points out several flaws in the previous studies. Foremost, he reminds that the MIMIC model assumes that the relations the indicator variables have with the causal variables are solely carried through the latent variable. In other words, the specifications used by the existing studies assume that there are no direct effects between the cause variables and the indicators - an assumption which is not likely to hold. It is, however, possible to identify some direct and indirect effects, and the identification and estimation of direct and indirect effects has been discussed by Fox (1980) and Bollen (1987).

### 2.3 Economic Freedom as a Driver for Growth in Transition

The quality of government is critical for economic success since the government affects the economy at least via two channels. First, the government has an indirect
role as an economic player, since it provides the economic and political institutions. Market reforms and controlling against opportunities for rent-seeking and corruption are emphasized (see Frye and Shleifer (1997) and La Porta et al. (1999)). Vested interests for government officials can halt the reform process, create incentives for the underground production, and slow progress in the official economy (see Harstad and Svensson, 2006). As Hay and Shleifer (1998) puts it, sometimes the failure of the state to enforce its own laws is due to weak incentives in the government to provide law and order. In the grabbing-hand model of Frye and Shleifer (1997), the government is above law and uses its power to extract rents. Policies and institutions are shaped by those in power to stay in power and to transfer resources to themselves (for example Turkmenistan under Saparmurat Nijazov’s rule). For example, Johnson et al. (2000) report a finding that for 90% of Russian and Ukrainian managers it is a custom to bribe the government officials, yet the firms in their industry also pay for ‘protection’ of their activities.

Second, since the government is an independent economic entity with its own income and spending, it may directly affect the economy altering the composition of its budget. Typically the public investment can spur the growth, while the taxation and wasteful spending may depress it. Clearly these two functions of the government interact and are not set independent on each other. How the government allocates its budget is largely dependent on whether the government is working for the public or private interest.

The bad economic policy and corruption are likely to promote the informal economy. At least part of that production will be away from the official economy hence the tax base, while the income created in informal economy will contribute to the economic well-being of citizens. According to Schneider’s (2004) estimates, the size of the shadow economy in transition countries ranged from 20.1% in Czech Republic to 68.0% in Georgia at 2002/2003. These omissions of income and intranational transfers are huge, whereby the welfare comparisons and the productivity growth estimates that account only for measured production (and observed transfers of income) are misleading. This idea has already received some attention, as Feige and Urban (2008) review the issue of measurement. They concentrate on the measurement of non-reported income and acknowledge that until the problems
of measurement are resolved, investigations of the relationship between economic reforms and economic outcomes during the transition decade must be viewed with considerable caution.

There are papers detecting the role of the institutions on the economic growth in the context of economies in transition. Fidrmuc and Tichit (2007) find that the pattern of growth in transition has changed at least twice; yielding three different models of growth associated with different stages of reform. In the last stage of transition, which is model three, the reform has a positive effect on growth.\(^5\) Iradian (2007) finds that growth in CIS (The Commonwealth of the Independent States) has been higher because of the recovery of lost output, progress in macroeconomic stabilization and market reforms, and favorable external conditions. About half of the growth is due to progress in macroeconomic stabilization and market reforms. The other sources of growth seem to be outside the government’s control. Lastly, Babetskii and Campos (2007) review the studies trying to investigate the effect of institutional reform to growth and conduct a meta-analysis on others’ results. They find that approximately one third finds positive and significant relationship, another third finds negative and significant relationship, while the final third find no significant relationship between reform and growth. In line with Fidrmuc and Tichit (2007), they also show that the estimated relationship between reform and growth seems to change over time. In addition, they note that existing results are sensitive to the choice of the measure of reform used in these studies. Bearing Feige’s and Urban’s (2008) criticism in mind, it is interesting to see how the reform policies affect growth of total output or GDP rather than the observed one.

The number of empirical, cross-country studies on growth is huge and this is also reflected in the existing number of models and variables. Durlauf et al. (2005) list 145 potential explanatory variables in growth regressions. Recently, Magnus et al. (2008) took an attempt to reduce the model uncertainty and to determine the ‘focus’ and ‘auxiliary’ regressors for growth. They find that constant, initial GDP per capita, real equipment investment share of GDP, initial total gross enrollment share of GDP, and many more.

\(^5\) Also Roland and Verdier (2003) find different kinds of trajectories of economic transition from socialism to capitalism, and the failures in law enforcement are not so bad in other transition economies as in Russia.
ratio for primary education and the initial life expectancy at age 0 are the focus variables. In other words, these variables should be included or at least tested in all empirical investigations. Minier (2007) tests the presence of fiscal variables and non-linearities of growth model and shows that either squared terms or interaction terms of fiscal variables should be present in growth models. Furthermore, allowing for the non-linearities the several fiscal variables become robust - a finding that is confirmed by some previous studies.

On the public expenditure, Munnell (1992) criticizes the proponents of spending advocates for their failure to recognize the common trends in public capital and productivity. Devarajan et al. (1996) find that the share of current expenditure has a positive effect on growth, while an increase in the capital component of public expenditure has a negative growth effect. Aschauer (2000) argues that the relationship between public capital and economic growth is non-linear; permanent changes in public capital are associated with permanent changes in economic growth. Fölster and Henrekson (2001) find a negative relationship between government expenditure and growth in rich countries, while Blankenau et al. (2007) find a positive relationship between public education expenditures and long-term growth after controlling for government budget constraint.

2.4 Economic Growth and Demographic Transition: A Club Approach

The literature on the cross-country convergence of incomes is full of controversies and puzzles. One of them is that convergence fails in heterogenous samples of countries if growth is non-linear in key factors. To control for the observed heterogeneity the theory and practice offers several factors: mutual trade (Ben-David and Loewy1998), common history (Baumol 1986), geographical location (Maddison 1994) etc. Recently, Galor (2007) and Galor and Weil (2000) among others, revisit the centuries old hypothesis postulated by Thomas Malthus (1798), whereby the population grows in geometrical ratio while the subsistence only increases in an arithmetical ratio. While Malthus suggested that if unchecked the population will grow beyond its subsistence the modern researchers present more optimistic views. Firstly, Galor (2007) suggests that an uneven demographic transition may explain
several controversial observations. Secondly, differences in the timing of the transition have segmented countries to different regimes, and they exist simultaneously. The process, however, is dynamic whereby all countries should progress through these regimes until they reach the stable growth regime.

In the demographic transition considerable changes take place in all demographic related factors, yet different authors have emphasized different perspectives. Soares (2007), for example, maintains that the infant mortality has a leading role; if it starts to fall fertility will follow, and the transition will proceed according to its own internal laws. Ram (1998), Fogel (1994, 2004), Becker et al. (2005), and Lorenzen et al. (2008) claim that life lengthening is the revolutionary factor, as short-sighted and deterministic attitudes give way to optimistic views. With regards to the number of demographic clubs or stages Bloom and Williamson (1998) stress the role of dependency rate and argue that the number of clubs should be limited to two, namely to “Demographic Burden” and “Demographic Gift”. Galor and Weil (2000), on the other hand, advocate the presence of three stages labeling them the “Malthusian”, ”Post Malthusian” and “Modern”. The first one is best described as a demographic trap in spirit of Malthus, while the second is characterized as a stage of industrial revolution or a take-off and the last stage is the regime of sustained growth.

The early empirical literature suggested that the OECD countries exhibit a tendency to convergence, while results for Africa, Latin America and Asia-Pacific were more mixed. Recently Li and Papell (1999) and Pedroni (2007) present evidence for the non-convergence of the output in the OECD, and suggest that only a fraction of countries exhibit tendency for convergence. With the exception of Pedroni (2007), all of these works suffer from an important drawback: the selection of countries is prone to the cross-section dependency biasing the test results. In the context of output convergence, this means that convergence can be found where it does not exist. To correct for this bias Pesaran (2007a) proposes a use of a modified panel unit root test in the presence of a single unobserved common factor.

Demographic transition is a dynamic process, far from completed in many countries, whereby in the future countries may move ahead toward more mature demographic stages. Thereby the demographic transition can increase our understanding
about the distribution of world incomes and foreshadow the prospects of convergence in the future. These insights have some relevance to policy-makers, not only for the sake of demographics, but for the sake of equalizing the distribution of world incomes. Uncovering the demographic dynamics and its impact on growth is therefore prerequisite for forming the right policies to tackle poverty.

3 Summary of the Essays

3.1 Optimal Law Enforcement in the Presence of Organized Crime

In this essay we build a model, to evaluate the interaction between the government and the mafia, and to study the law enforcement of underground production in the presence of organized crime. The question asked, is how the size of the shadow economy, the optimal tax and public goods are affected by the policing of the shadow economy in the context of dys-functional government. We evaluate how the government and firms are affected by the presence of organized crime. We ask, who gains and who loses when the illegal producers can be monitored and punished by the government that is not concerned on the welfare of the citizens.

We analyze the interaction between the government and the crime group (aka the mafia) in a context of a relatively stable state. We consider a sector of economy where the business can be plied either legally or illegally and where the firm makes this choice in the presence of government’s and mafia’s rent-seeking. The government is not concerned on the welfare of the citizens; instead it maximizes its own payoff using taxation, monitoring and public goods as a tool. The illegal firms have a partial access to public goods. We assume that the government does not have enough resources to fight the organized crime, but can monitor the illegal production.

We first solve the model in the absence of the mafia as a benchmark. We then introduce the organized crime and explore both the static and the dynamic games. To compare our results with previous studies, and to name the winners and the losers, we consider the game without monitoring.
The main results are the following. Firstly, it is always the case that the revenue-maximizing government will monitor the shadow economy, with and without the presence of mafia, as monitoring contributes to the government’s revenue. Secondly, the government’s tax policy is backed up by its monitoring policy that is the rate of tax is increased by the expected punishment at the shadow sector. Thereby the government uses its power to monitor to extract more from the *formal* economy. Thirdly, the entry of the mafia increases the government’s revenue since the exit option is now more expensive. By the same token, both legal and illegal firms are, in general, hurt by the entry of the mafia. It is, however, plausible that legal firms might benefit by the entry of the mafia.

Results related to monitoring give some additional insights. Firstly, the size of the shadow economy is independent of monitoring. Second, the legal and illegal firms are better off without monitoring, whereas the government is worse off. The mafia goes completely unaffected by monitoring. From the normative point of view, whether the government should have a possibility to monitor shadow economy depends on whose benefits are emphasized. If the fine is viewed as a hidden tax or a bribe, then corrupted governments should be stripped off their right to punish offenders.

### 3.2 Do The Catholic and Protestant Countries Differ by Their Tax Morale?

In the second essay, the key question is whether a religious denomination can explain differences in tax morale. For example, in the southern catholic countries, religion has a built-in tradition of forgiveness. One is tempted to suggest that this may support an equilibrium where the social punishment of, say, tax evasion, can have a different role than in the protestant countries. Should this hypothesis be valid, it will be reflected in the elasticity of the size of the shadow economy with regard to tax morale. To evaluate the effects of tax morale, taxation, government consumption, social security, crime and interest rates on the shadow economy we work with the data from the OECD. We estimate both the direct and indirect effects between some causes and indicators.

In our data, we identify two regimes. The first regime, 1979-1992, can be char-
acterized as a regime of public sector growth, whereas during the latter regime, 1992-2003, the growth of the public sector came to an end. To test the hypothesis that religion affects tax morale and hence tax evasion, we use the data from the World Values Surveys to group the countries according to their primary religious denomination. The countries form natural groups of the catholic south and the protestant north, whereby the other countries make up the reference group that facilitates the comparison between groups and with the previous studies.

We report three major findings. First, there are some differences in tax morale across religions, but for the latter regime we reject the hypothesis that countries with a different religious background exhibit different tax morale. Our estimates suggest that within the OECD, the norms have become global. Second, even though we do not calibrate the index of the size of the shadow economy, our results suggest that the shadow markets are probably not of the magnitude that has previously been proposed. The main reason for this result is that we have included the direct effects of the cause variables on the indicators. Third, our results support the view that there was a regime switch in the evolution of the public sector in the early 1990s and that the evolution of the shadow economies within these regimes differs.

3.3 Economic Freedom as a Driver for Growth in Transition

The evidence for the idea that good governance is critical for economic success is unequivocal, creating a need to verify whether this view is true in transition economies. As the informal production makes up a big portion of total production in these countries, we take an attempt to correct for this bias by using real GDP per worker. By concentrating on those who are listed as to work in formal economy, we better measure the average productivity and the growth potential of these nations.

There are several findings in this paper as to the economic growth in transition. First, we find no regimes on growth; all countries surveyed obey the same model and laws of motion and the relationship between growth and human capital (institutions) is linear across space. Second, at the beginning economic freedom and investments contribute positively on growth in transition. When the countries reach their population averages in terms of economic freedom or investments the negative interaction takes its toll; the full marginal effect of investments (economic
freedom) to growth is negative if the economic freedom is above its average. As economic freedom (investments) increases, greater investment (economic freedom) appear to have a detrimental impact on growth, suggesting a joint “too much of a good thing” phenomenon in transition economies. Third, contrary to some earlier findings, our results suggest growth is influenced by government policies on shaping institutions and setting the size of the government. This influence is, however, more complicated than previously thought.

All these findings are robust when we drop resource-rich countries from the dataset. We show that non-linearities are present in the growth model in terms of interactions. Our results indicate that the somewhat contradictory results in the earlier literature might be due to use of an inappropriate model. Moreover, we find that how we measure institutions makes a difference; our results change when we use Fidrmuc and Tichit’s recalibrated index of economic reform. Finally, our robustness analysis shows similar results when evaluating real growth per capita, although investment and size of government seem to matter more for productivity growth than growth per capita.

A useful observation is that growth researchers should use care in selecting indicators for measurement of economic well-being, institutions and the model. The main policy implication is that some transition countries might have over-done both private investments and marketization, which has hurt the productivity growth.

3.4 Economic Growth and Demographic Transition: A Club Approach

On the relationship between economic growth and demographic transition, our main findings are as follows. We find that the non-linearities are present. The number of the demographic clubs is three and the discriminating variable turns out to be life expectancy, since it gives incentives and possibilities to accumulate human capital and to invest. Furthermore, the increasing life-expectancy status from Club I to Club III seems to give support to their interpretation as three successive demographic stages, as proposed by Galor and Weil (2000).

We find $\beta$-convergence in all clubs and divergence in the full sample, i.e., our club approach has succeeded in uncovering three traditional convergence clubs on
the demographic basis alone. As the $\beta$-convergence, however, is probably not valid, we used both first and second generation unit root tests as well.

For the club of low-life expectancy some of the traditional tests support conditional convergence with fixed effects and individual trends. Moreover, there is also some support for the typical finding of the (un-) conditional convergence among the high-life expectancy club. The more robust second generation unit root test undermines this result. This finding is in line with the trends in the literature, where the earlier tests suggest convergence but the most recent tests show that homogenous clubs are difficult to uncover, even among the industrial countries. The demographic transition is a continuous rather than discrete process, leaving some heterogeneity in the timing of it even within the clubs. Furthermore, the phase of transition differs across countries. Over time, this gives rise to the heterogeneity within the clubs that cannot be controlled for simply by controlling the initial life expectancy. These concerns regard Club II most seriously, in which take-offs and leapfrogging not only in terms of output but also in terms of life expectancy is typical. In this club our results show no convergence by any of the tests. The rapid progress in the field of panel estimation techniques may shed further light on this subject in the future.

Lastly, as the demographic dynamics seem to support transitional rather than permanent clubs, we forecast the future of these countries and extrapolate their real GDP per capita by using Club-specific growth rates. This information gives no support to the convergence of world incomes; the income gaps have already widened to such an extent that even the take-off is unable to rise the incomes of the poor sufficiently. Thus, new economic miracles will hardly arise on the demographic basis alone. Therefore, economic policies should be targeted to help that minority which is left behind in the lowest club. With regard to economic policy, an interesting question is, whether the most efficient policy is to concentrate on health and life-expectancy, as longer and healthier life should give incentives to save and invest in human capital, both of which are the impetus of growth.
References


Plato (360 B.C.E.): Republic.


Optimal Law Enforcement and Welfare in the Presence of Organized Crime

November 12, 2009

Contents

1 Introduction

2 The model

3 The birth of shadow markets in the absence of a mafia

4 The mafia arrives

4.1 Simultaneous moves

4.2 Stackelberg game

5 Welfare and enforcement

6 Conclusions

A Restrictions for interior solutions

A.1 Sub-rules for the existence

A.2 The regions

B Proofs

B.1 Cournot vs. No mafia

B.2 Cournot vs. Stackelberg

24
1 Introduction

Many countries host crime groups (mafias, clans, triads, gangs, etc.) that engage in illegitimate activities such as drug dealing and extortion. Certain government actions (poorly designed tax policies, excessive or heavy-handed regulation, failures to deal with corruption, etc.) can distort the market equilibrium and create incentives for firms to enter the shadow economy or black market.\(^1\) This shifts the fiscal burden onto honest actors. However, legitimate firms, unlike their counterparts operating in the shadows, often enjoy full access to goods and services produced by the government. On the other hand, the government can inadvertently, negligently, or even deliberately, create rent-seeking opportunities for nefarious operators. In the worst case, citizens and businessmen find themselves vulnerable to extortion from both the government and the mafia (Johnson et al., 2000 and Los, 2003).

Along these lines, this paper considers the economic consequences of a Leviathan government that monitors and punishes underground production. This issue is particularly relevant for Europe’s transition economies, which are still experiencing major institutional reforms as they move to market economies. A common pattern seen in European transition experiences is that the state enters the process enfeebled, resulting in a boom in illicit activities and organized crime (e.g. Johnson and Kaufmann, 2001). As the transition process advances, however, reform of the state apparatus and institutions forces decision-makers to seek optimal policies on taxation, provision of public goods, and the judiciary.\(^2\) At the start of transition, proponents of the “big bang” approach argued that the market economy offered such a huge benefits so countries could initially forego efforts at democratizing the political system and focus on reform of the economy. Nearly two decades into this process, we see transition economies where democratization has yet to materialize

---

1Loyaza (1996) observes the size of the shadow economy is positively related to the tax burden and negatively to the quality of government institutions. Johnson and Kaufmann (2001) find the underground economy in transition economies has been driven mainly by excess regulation, corruption, a weak legal environment, and, to a lesser extent, by taxation.

2Torgler (2003) notes that it has been difficult to find the right equilibrium of state activity in transition economies; there are strong traditions of state interventionism and bureaucracy without adequate protections for property rights. In the early years of transition, post-communist governments often extracted rents for their own use. Eventually, many governments evolved into constrained Leviathans, i.e. non-welfarist, revenue-maximizing states that produce a few public goods to keep people happy. See also Konford (2000).
despite their big banging. Instead, we find Leviathan-type governments occupied by members of the *nomenklatura* and persistent failures in the market and political sphere (e.g. Åslund, 2007 Ch 9.).

Previous work has considered the origins of organized crime and the interaction of the government and a mafia. Anderson (1995) considers three examples of conditions that fostered the emergence of mafias: the abdication of legitimate government power (Sicily), excessive bureaucratic power (former Soviet Union), and the potential of illegal markets (United States). Williams and Godson (2002) propose that a weak state, lacking in democracy or rule of law, offers fertile soil for the growth of organized crime. In the “grabbing-hand” model of Frye and Shleifer (1997), a weakened Leviathan-type state is characterized by predation regulations, corrupt officials and a dysfunctional legal system.3

Several studies take up government-mafia interaction, particularly in the context of law enforcement. Anderson and Bandiera (2000) study mafias as enforcement coalitions to protect property from predators.4 Grossman (1995) builds a model where public services enter in the production functions of legal and illegal firms alike, while the mafia produces a public good solely for illegal firms. As long as the government remains viable, the lot of the private producer improves as competition between the government and mafia increases the production of public goods. Alexeev et al. (2004) emphasize the role of the mafia as a provider of public goods. They show that when public goods are unimportant, the government actually benefits from the presence of the mafia as the mafia makes underground production costlier.

Based on what follows, it is not clear how the shadow economy or the mafia affects the society as a whole. The social welfare might be inadequate measure for the total welfare as for instance, it does not account for the fairness (see e.g. Kaplow and Shavell, 2002). The standard is to account both the costs and benefits from crime to the social welfare and deviations from the tradition are hard to justify. Along these lines, some conclusions regards to the desirability of shadow economy and organized crime have been offered, although the treatment of social harm varies.

---

3Skaperdas and Syropoulos (1995) model the emergence of gangs in the absence of control by the state. Anderson and Bandiera (2000) point out that even strong, welfarist states like Japan host mafias. Moreover, they show that wherever a mafia has emerged, it has been difficult to root out.

4A popular view is that Italy’s Cosa Nostra emerged as a defense mechanism for impoverished rural peasants against their landlords (e.g. Allum and Sands, 2004).
Garoupa (2000) and Mehlum et al. (2003) treat crime groups as rent-extracting, harmful organizations that do not offer public goods. Garoupa shows that as long as extortion constitutes a costless transfer from individuals to the criminal organization, extortion improves the social welfare by decreasing the incentive for harmful criminal offenses and allowing the government to reduce spending on law enforcement. Mansour et al. (2006) treat gang structure as endogenous; an increase in deterrence can lead to an increase in the number of competing criminal gangs and raise total illegal output. Bowles and Garoupa (1997), Garoupa and Jellal (2007), and Kugler et al. (2005) evaluate the case where the police and criminals collude. The findings of all these researchers contradict Becker’s (1968) famous maximal fine result that proposes that increasing policing and sanctions beyond an optimal point merely boost crime rates.

The paper studies a question that is completely new, i.e. how might policing the shadow economy of a rotten state affect the size of the shadow economy, taxation policy, and production of public goods. We evaluate how organized crime influences the government and firms. In particular, we consider who gains and who loses when illegal producers are monitored and punished.

We analyze the interaction of the government and the crime group in a context of a relatively stable state. As in Alexeev et al. (2004), we consider a sector of economy where a business can be conducted either legally or illegally, e.g. construction, trash-hauling and harbor services. The firm makes a choice between legal and illegal production in the presence of government and mafia rent-seeking. Following the argument put forth by Baumol (1995), the Leviathan government is not concerned with the welfare of its citizens but rather maximizing its own payoff. The Leviathan

---

5 Berkowitz and Li (2000) study a situation common in transition economies: poorly defined government tax rights. As a result, other agencies levy their own taxes on the same tax base. A standard tragedy-of-the-commons problem emerges in which the tax base is “over-grazed.” The economy faces two equilibria according to the number of tax agencies with the implications for the share of illegal production, government’s tax revenues, and production of public goods.

6 Konrad and Skaperdas (1998) reach the same conclusions with a model in which the gang extorts from both legal and illegal enterprises.

7 See Garoupa (2007) for a similar analysis.


9 Baumol (1995) comments that dynasties are typically established through violence; governments concerned with the welfare of citizens and constrained by the law are historical rarities. See also La Porta et al. (1999).
state produces public goods to which illegal firms have partial access. The crime
group does not provide protection or anything else in exchange for the rent, but
aims to maximize its rent income in accordance with the cost of collecting rent.
We assume the government lacks adequate resources to fight organized crime, but
can monitor illegal production. No collusion of government agents and criminals is
allowed. Since the act of illegal firm itself is legal, but the firm is extralegal in the
sense that it either has no operating license or it evades taxes, we assume there is
no direct cost from the existence of an illegal firm. The cost arises in terms of fiscal
externality and the possibility that the provision of public goods is reduced.

We first solve the model in the absence of the mafia as a benchmark, then
introduce organized crime and explore both the static and the dynamic games. To
compare our results with previous studies and name the winners and losers, we
consider the game without monitoring.

The main results are the following. First, it is always the case that a revenue-
maximizing government will monitor the shadow economy, regardless of the presence
of a mafia. Second, a policy of monitoring supports government tax policy. Tax
rates can be increased if the shadow sector expects punishment for tax evasion, i.e.
the government uses its power to monitor to extract more from the legal sector.
The entry of a mafia increases the government’s revenue since the exit option of
a legal firm to the illegal sector is now more expensive. By the same token, both
legal and illegal firms are hurt by the entry of a mafia. There are, however, subsets
of solutions where illegal and legal firms, the latter in particular, can potentially
benefit from the entry of a mafia. In one case, two evenly matched “authorities”
compete in the Cournot competition, causing the government to decrease the tax
and increase the production of public goods to downsize the fiscal burden. This
option is not available in a Stackelberg competition, which is the worst case for
both types of firms.

Results related to monitoring provide additional insights. The size of the shadow
economy is independent of monitoring, because the government increases taxes when
it monitors, while the public goods are unaffected. Moreover, legal and illegal firms
alike benefit in the absence of monitoring, while the government is the clear loser.
The mafia goes completely unaffected by monitoring. From the normative point of
view, whether the government should monitor shadow economy depends on whose
benefits are emphasized.

The paper is structured as follows. Section 2 phrases the model and Section
3 presents the benchmark results absent organized crime. Section 4 introduces a
crime group into a Cournot game and a Stackelberg game. Section 5 discusses the welfare effects of monitoring under various assumptions. Section 6 concludes.

2 The model

Assume a continuum of firms with the mass of one and that each firm produces one unit of output at the market price of one. Following Alexeev et al. (2004), a firm entering an industry faces a choice between legal or illegal production. The firm chooses its sector by comparing the potential payoffs or refrains from production if it cannot make the reservation profit 0. Although legal production is subject to tax \( t \), the firm benefits from public goods \( g \) produced by the government. We write the payoff \( \pi_l \) of a legal firm as

\[
\pi_l = 1 - (t - g),
\]

where \((t - g)\) is the net fiscal burden borne by the legal firm.

An illegal firm, in contrast, pays rent \( r \) to the mafia.\(^{10}\) The public goods are partially excludable and parameter \( b \) represents the factor of public goods enjoyed by the illegal firms \( 0 < b < 1 \). An example of a public good denied from illegal firms would be contract enforcement. On the other hand, the illegal firm clearly benefits from many public goods such as roads and other infrastructure. Illegal firm also faces an exogenous penalty, \( z < \bar{z} \), if their activities are discovered by government monitoring agents. We assume that the fine is not optimized but instead it is set by an independent court, which is yet to be reformed. The probability of getting caught is \( p \).\(^{11}\) Thus, payoff \( \pi_{ij} \) for the illegal firm may be described as

\[
\pi_{ij} = 1 - r + bg - pz - c_j,
\]

where \( c_j \) is the cost of operating in shadow economy for firm \( j \). By setting (1) equal to (2), we solve the threshold cost for going underground

\[
\bar{c} = t + (b - 1)g - pz - r.
\]

\(^{10}\)Here it is assumed that the rent is the only cost the firms face because of the mafia. It could well be that both legal and illegal firm are harmed by a cost, say, \( k \), which would not affect the outcome as long as both types of firms must deal with it and the mafia draws no benefit from it.

\(^{11}\)One way to see the role of \( p \) and \( g \) are to think of them as the institutions provided by the government.
Firms with $c_j > \bar{c}$ enter the legal sector; the rest enter the illegal sector. The cost of operating in the shadow economy, $c_j$, can be viewed as the cost difference between legal and illegal production, where the negative values suggest that some firms benefit from going underground. Extra costs associated with illegal production include the lack of access to the formal credit and finance. Benefits may include freedom from accounting systems and operating licenses. Indeed, much of the bureaucratic burden facing legal firms is avoided by going underground. Whether the costs or benefits dominate depends on the skills of the entrepreneur, i.e. the sign of $c_j$. The cost is distributed uniformly between $[-x, x]$, where $x < 1$ illustrates the degree of the heterogeneity of entrepreneurs; the larger the value of $x$, the more heterogeneous community of entrepreneurs. The cumulative distribution function of $c$ and the size of the shadow economy is

$$F(c_j) = \frac{1}{2x} (\bar{c} + x), \quad (4)$$

where $\bar{c}$ is defined by (3).

The government maximizes its profits that are the revenue from taxes and from fines minus the production of public goods and monitoring services. We simplify the model by assuming that collecting taxes involves no costs. Next, the government chooses $t, g$ and $p$ to maximize its payoff

$$G(t, g, p) = (1 - F(\bar{c})) t + F(\bar{c}) pz - \frac{1}{2} g^2 - \frac{1}{2} p^2, \quad (5)$$

subject to the constraint that $\pi_l(\bar{c}) \geq 0$, i.e. the marginal firm will receive its reservation profits.$^{12}$ In the equation (5), the first term is the expected tax revenue from the legal sector, the second term is expected income from fines from the illegal sector and the last two term are the cost of producing public goods and policing.

Similarly, the payoff for the mafia is defined as the income from rent-extracting minus costs of collecting the rent which is relative to the size of shadow sector. More firms there are in the shadow economy the higher the cost of collecting protection fees. Thus, the mafia will seek to maximize its payoff with respect to $r$

$$M(r) = F(\bar{c}) r - \frac{a}{2} F(\bar{c})^2. \quad (6)$$

where $a > 0$. The timing of these games is such that in a static game, the government and the mafia move simultaneously and are followed by the firms. In a

$^{12}$Here, all firms choosing to remain in the legal sector receive the same profit as the marginal firm. Those in the illegal sector obtain higher profits due to their entrepreneurial skills.
dynamic game, the legal authority, i.e. the government moves first, then the extra-
legal authority, i.e. the mafia. Only after the mafia has moved do firms decide
whether to move themselves.

In the present analysis, we focus on interior solutions and the comparability of
results among different scenarios (a detailed description is given in Appendix A). To
this end, we make certain restrictions on $x$ and $z$. As stated, it is a reasonable intu-
ition that the cost of operating underground and the penalty of discovery are both
less than the income generated by illegal production. Obviously, rules governing
legal production, i.e. $\pi_l(\bar{c}) > 0$, must exist in all these set-ups.

To make our three set-ups comparable and obtain interior solutions, we must
identify the sub-set of $\{x, b, z\}$ where all these solutions exist. Assume that $3x >
(b - 1)^2$ is true. Then we write Assumptions 1 to 2:

1. Assume $2x > (1 - b)$ and $z^2 < 1 - \frac{(4x+1)+(b-1)(3x+1)}{2(4x+1)-(b-1)^2}$ are both true (Region
1).

2a. Assume $2x < (1 - b)$ and $z^2 < 1 - \frac{(4x+1)+(b-1)(3x+1)}{2(4x+1)-(b-1)^2}$ are both true (Region
2a).

2b. Assume $2x < (1 - b)$, $4x < (1 - b)$ and $z^2 < 1 - \frac{(4x+1)+(b-1)(3x+1)}{2(4x+1)-(b-1)^2}$ are all
true (Region 2b).

In the essence, the restrictions here are relations between the heterogeneity of
the firms and their ability to benefit from the public goods in each sector. $(1 - b)$
denotes the extra benefit from the public good if the firm stays at legal sector,
while $x$ is the highest possible expected extra cost that the firm faces if it goes
underground. In Region 1, if the extra costs of going underground are high com-
pared to the extra benefits of remaining in the legal sector, then the government
should not collect an excessively high penalty to keep both sectors viable in all these
set-ups. With less heterogeneity, the need to set restrictions to the penalty arises,
particularly in the case of the Stackelberg game.

Figure 1 illustrates regions where different outcomes exist. Limits to the penalty
are not considered here, but are expected to hold. Bold curve illustrates restriction
$3x > (b - 1)^2$, i.e. the solutions exist only above it. The line farthest to the left
illustrates restriction $2x > (1 - b)$, above which we operate under Assumption 1 in
Region 1. If we operate below $2x < (1 - b)$ but above $3x > (b - 1)^2$, we are either
in Region 2a or 2b (Assumptions 2a 2b). Outcomes that are true in Region 1 are
not necessarily true in Region 2a and Region 2b.
Figure 1: The bold curvature illustrates restriction $3x > (b - 1)^2$ while the straight line furthest to left illustrates restriction $2x > (1 - b)$. The shaded region in the upper-left corner is Region 1, while Regions 2a and 2b are areas below it, Region 2b being below Region 2a.

Equations (1) to (6) with Assumptions 1 to 2b describe the full model, which includes shadow production, a mafia, and the possibility to monitor by the government. Most of the following results are true under all assumptions; thereby only exceptions for this rule will be indicated. First, we address our baseline scenario: a mafia-free economy.

3 The birth of shadow markets in the absence of a mafia

Obviously, if there is no shadow economy, firms have nowhere to go to escape the grabbing hand of the government. The government, in turn, has full power to determine the tax rate, what public goods it provides, and appropriate a part of the profits of firms. Of course, if a firm is left with nothing, entrepreneurs have a strong incentive to start producing illegally.

First, we rewrite equation (2) as

$$\pi_{ij} = 1 + bg - pz - c_j.$$ 

The threshold cost for firms entering the shadow economy is obtained from the
indifference \( \pi_t = \pi_{ij} \), with the result that \( \bar{c}_1 = t + (b - 1) g - p z \), where subscript 1 indicates the game in the absence of the mafia, i.e. \( r = 0 \). Thus, the size of the shadow economy, i.e.

\[
F(\bar{c}_1) = \frac{1}{2x} (\bar{c}_1 + x) = \frac{1}{2x} (t + (b - 1) g - p z + x),
\]

increases as taxation rises or public goods, penalties or monitoring are reduced. It is crucial here to acknowledge that there are indirect costs transmitted via changes in the size of the shadow economy in addition to the direct costs of producing public goods and monitoring.

The payoff for the government follows equation (5). Subject to \( \pi_t (\bar{c}) \geq 0 \), we optimize the Lagrangian with respect to the tax rate, public goods, and monitoring. The FOCs are

\[
\begin{align*}
\frac{\partial G}{\partial t} = 0 & \iff \frac{\partial F(\bar{c}_1)}{\partial t} (p z - t) + (1 - F(\bar{c}_1)) = 0 \\
\frac{\partial G}{\partial g} = 0 & \iff \frac{\partial F(\bar{c}_1)}{\partial g} (p z - t) - g = 0 \\
\frac{\partial G}{\partial p} = 0 & \iff \frac{\partial F(\bar{c}_1)}{\partial p} (p z - t) - p + F(\bar{c}_1) z = 0,
\end{align*}
\]

On the basis of (8), we put forth the following proposition:

**Proposition 1** The government should strive to collect taxes at an optimum rate \( t > 0 \). Moreover, monitoring should always be part of an optimal policy from the government’s standpoint.

**Proof.** Since \( 0 < F(\bar{c}_1) < 1 \), it must be that \( \frac{\partial G}{\partial t} \big|_{t=0} = (1 - F(\bar{c}_1)) + \frac{\partial F(\bar{c}_1)}{\partial t} p z > 0 \) for all non-negative \( p \). Similarly, setting \( p = 0 \) gives \( \frac{\partial G}{\partial p} \big|_{p=0} = - \frac{\partial F(\bar{c}_1)}{\partial p} t + F(\bar{c}_1) z = 0 \).

The intuition is that the government does not incur direct costs from taxation even with indirect costs from an increase in the size of the shadow economy. Clearly, when \( t \) is small, benefits exceed the costs of taxation. Note that when \( p = g = 0 \), the size of the shadow economy in equation (7) is never zero as long as \( t > 0 \). The intuition for the second result is that when \( p \) is small, income generated under a monitoring regime exceeds the costs of monitoring. Therefore, monitoring increases the payoff for the government. The optimal solution for the tax rate, the amount
of public goods and the monitoring are

\[ t_1 = \frac{-2x^2}{-4x + (b - 1)^2} + z^2, \]
\[ g_1 = \frac{x(b - 1)}{-4x + (b - 1)^2}, \]
\[ p_1 = z. \]  

(9)

We now analyze how exogenous variables affect the optimum, noting first that increases in \( b \) decrease both the tax rate and the amount of public goods. The intuition here is that when \( b \) (the benefit an illegal firm draws from public goods) increases, production in the shadow economy becomes more attractive. The government reacts by decreasing the production of public goods (which now hurts the illegal firm more than previously) to decrease production costs. At the same time, decreased taxation makes operating in the legal economy more attractive. Indeed, the increased size of the legal sector may be enough to offset the effect of lost government income from an increased \( b \).

Second, increases in the penalty \( z \) always increase the tax rate and monitoring without affecting production of public goods. In particular, since \( p_1 \cdot z \) is \( z^2 \), the firm faces a penalty in one form or another, and that cost will not affect the choice of sector. This comports with expectations. From (5), we observe that when the government is able to monitor, albeit the cost for the illegal firm is probabilistic, so the government can collect at least some rent from all the firms in the economy. From (9), we observe that we may write \( t_1 = p_1 z + d \). Setting the expected penalty as a minimum rent collected from all firms, while collecting something extra from legal firms, say \( a \), the income for the government can now be rewritten as \((1 - F(\bar{c})) (p_1 z + d) + F(\bar{c}) p_1 z = p_1 z + (1 - F(\bar{c})) d\). Thus, income \( p_1 z \) is always certain for the government, and since firms will face it at both sectors it will not affect firm’s decision on which market to enter. The size of the shadow economy can be written as \( \frac{1}{z^2} (d + (b - 1) g + x) \), i.e. the effect of monitoring on the shadow economy is completely neutralized. Therefore, increases in the penalty \( z \) lead, one to one, to increases in monitoring, while the government neutralizes the effect of the expected penalty by increasing one to one the tax rate and leaving everything else unaffected. In essence, monitoring not only increases the (fiscal) burden borne by the illegal firm but also the fiscal burden borne by the legal firm. Moreover, since public goods are not affected by monitoring, the legal firm is still the sole contributor to the production of public goods, while the income from monitoring is
wasted.

Finally, increases in the heterogeneity of the skills of entrepreneurs, \( x \), decreases the amount of public goods and increase taxes. Monitoring remains unaffected. Thus, the most skillful entrepreneurs can make more profit operating in the shadow economy, while those less skilled benefit more from remaining in the legal sector. The overall productivity of the firms increases, allowing the government to grab more revenue. Although increasing the tax rate and decreasing the production of public goods drives some entrepreneurs to the illegal sector, the government collects more from individual firms remaining in the legal sector and enjoys lower costs for production of public goods.

Substituting the solutions to the payoff for legal firms gives a strictly positive result; the government can never extract all revenues of legal firms:

\[
\pi_{ll} = 1 - z^2 + \frac{(b - 1) + 2x}{-4x + (b - 1)^2} > 0. 
\]

(10)

Evaluating (10), we find the revenue of legal firm decreases when \( b \) rises. An increase in \( b \) decreases the amount of public goods and the tax rate, so the net effect on the legal firm’s payoff is negative. Increases in penalty \( z \) decrease the revenue of the legal firm. Thus, we intuit that when the amount of public goods produced remains unchanged while the tax rate is increased with the penalty the net effect is to hurt the legal firm’s profits.

Solving for this threshold, we obtain

\[
\bar{c}_1 = \frac{-2x + (b - 1)^2}{-4x + (b - 1)^2}.
\]

Evaluating the threshold, we find that \( \bar{c}_1 < 0 \) if \( x < \frac{1}{2} (b - 1)^2 \), and positive otherwise. Since the penalty does not affect the threshold, we infer that the size of the shadow economy is also unaffected.

**Proposition 2** The size of the shadow economy is unaffected by monitoring.

Firms with a cost of entering the shadow economy below \( \bar{c}_1 \) will choose illegal production. The size of the shadow economy decreases if the benefit from public goods decreases for firms in the illegal sector. The higher the excludability of public goods, the greater the incentive for firms to stay in the legal sector. Increases in \( x \) boost the size of the shadow economy, giving the most skillful entrepreneurs more opportunities in the shadow economy.
Solving for the government’s payoff, we get

\[ G_1 = \frac{1}{2} \left( z^2 + \frac{x^2}{4x - (b - 1)^2} \right). \]  

(11)

The benefit the illegal firm draws from public goods decreases government’s payoff, while the government’s payoff is increased in the penalty. The ability to monitor benefits the government in two ways: fines increase direct income from the shadow economy and higher taxes increase income from the legal sector without disturbing the amount of public goods produced. In other words, monitoring supports government tax policy.

4 The mafia arrives

We now assume the market is occupied by two authorities, both capable of extracting rents. The government has two unique features: it can produce public goods, and, as sole legal authority, can monitor the shadow economy. On the other hand, the government is too weak to control the mafia fully or eradicate it. The choice a firm faces under this scenario requires comparing equations (1) and (2). The threshold cost for going underground is determined, following equation (3), by the decisions of the government and the mafia. For example, a decrease in rent \( r \), ceteris paribus, means more firms will find it profitable to operate underground. The size of the shadow economy is determined according to equation (4), and the government chooses \( t, g \), and \( p \) to maximize its payoff at equation (5). The mafia maximizes its payoff from equation (6) with regard to \( r \).

Now the game can be played either as a Cournot competition, where the mafia and the government move simultaneously, or as dynamic game as Stackelberg game where one party moves first. While the Cournot game is simpler, the Stackelberg game seems closer to reality as there is a likely pattern whereby the legal authority emerges first and the second authority (mafia) is extralegal by definition and follows in reaction.\(^{13}\) We first analyze the Cournot competition.

\(^{13}\)Skaperdas and Syropoulos (1995) view the gangs as primitive states emerging from anarchy. Since the legality of authority is important here, we expect that the first authority is legal and the second one extralegal by definition. Following Skaperdas and Syropoulos (1995) we note that the relationship between state and organized crime is seldom completely antagonistic and as suggested by Baumol (1995) we assume the government has intrinsic motives.
4.1 Simultaneous moves

The payoffs for the government and for the mafia in the static game are presented by equations (5) and (6), now denoted by subscript 2. To solve for the tax rate, monitoring, public goods, and the rent, we attempt to identify the best-response functions for the government and the mafia. Optimizing $G$ with respect to $t$, $g$, and $p$, results in the reaction functions for the government. Optimizing $M$ with respect to $r$ yields the reaction function for the mafia. From the FOCs of the government, we get

\[
\begin{align*}
\frac{\partial G}{\partial t} &= \frac{\partial F(\tilde{c}_2)}{\partial t} (pz - t) + (1 - F(\tilde{c}_2)) = 0 \\
\frac{\partial G}{\partial g} &= \frac{\partial F(\tilde{c}_2)}{\partial g} (pz - t) - g = 0 \\
\frac{\partial G}{\partial p} &= \frac{\partial F(\tilde{c}_2)}{\partial p} (pz - t) + F(\tilde{c}_2) z - p = 0.
\end{align*}
\]

(12)

Upon evaluation, we see Proposition (1) also applies in this set-up. The FOC for the mafia can be written as

\[
\frac{\partial M}{\partial r} = \frac{F(\tilde{c}_2)}{\partial r} r + F(\tilde{c}_2) - aF(\tilde{c}_2) \frac{\partial F(\tilde{c}_2)}{\partial r} = 0.
\]

(13)

Equation (13) illustrates how the optimal rent depends on the existence and size of the shadow economy. As $t$, $g$ and $p$ are not directly present in (13), government policies can influence mafia’s decisions only through the impact of government actions on shadow economy.

**Proposition 3** When there is a shadow economy, the mafia should strive to collect rents.

**Proof.** Evaluating the mafia’s FOC at $r = 0$ gives $\frac{\partial M}{\partial r} |_{r=0} = F(\tilde{c}_2) - aF(\tilde{c}_2) \frac{\partial F(\tilde{c}_2)}{\partial r} > 0$, since $\frac{\partial F(\tilde{c}_2)}{\partial r} < 0$ and $0 < F(\tilde{c}_2) < 1$. ■

Proposition 3 states that collecting rents is predicated on the existence of a shadow economy. But as was shown in Proposition 1, it makes sense for the government to monitor the shadow economy and try to extract rents when the activities of illegal firms are detected.
Solving the unknowns from these linear equations, we obtain

\[
\begin{align*}
t_2 &= \frac{2x(1 + 3x)}{(6x + 1) - (b - 1)^2} + z^2, \\
g_2 &= -\frac{(b - 1)(1 + 3x)}{(6x + 1) - (b - 1)^2}, \\
p_2 &= z, \\
r_2 &= \frac{(2x + 1)(3x - (b - 1)^2)}{(6x + 1) - (b - 1)^2}. \\
\end{align*}
\] (14)

These results are quite similar to the situation in a mafia-free system.\(^{14}\) Evaluating the mafia’s rent according to Assumption 1, we find it increases in line with benefit gained from partaking in public goods. Thus, the mafia “taxes” the increased benefit enjoyed by the illegal firm. Note that mafia rent is clearly below the tax rate, even in the absence of monitoring. Using the solutions in (14), we find the size of the shadow economy is diminished by the presence of a mafia (see Appendix B).

Comparing our results for with and without a mafia, we find that the optimal amounts of monitoring are the same. Evaluating the optimal tax and amount of public goods shows that \(t_2 > t_1\) and that \(g_2 > g_1\), i.e. while the emergence of a mafia allows the government to raise taxes, it also requires to increase production of public goods. How these changes affect the profits of firms’ depends in which region we operate. Operating under Assumption 1, as the government is unable to extract all profits of legal firms, we reach

**Proposition 4** *Both legal and illegal firms are worse off once a mafia enters the shadow economy.*

**Proof.** See Appendix B. ■

In Region 1, the net fiscal burden of legal firm is increased, while the net fiscal burden of illegal firm is decreased. The illegal firm, however, faces an extra cost in terms of rents, whereby its net profits are decreased and both firms are worse off. As soon as we exit Region 1 this is no longer the case. Instead, we reach,

\(^{14}\)In fact, assuming that also the legal firm had to pay rents for the mafia would lead the government making its decision as if there was no mafia. Then the results would be similar to those of the first set-up, with the exception that mafia’s payoff would increase and legal firm’s profits would decrease.
Proposition 5  The entry of a mafia benefits legal firms. For small enough $x$ (Assumption 2b) also illegal firms benefit.

In Region 1, the arrival of the mafia places the government in a stronger position vis-à-vis legal firms; it can increase taxes more than it increases public goods. While these changes create the incentive to go underground, the entry of a mafia increases costs of illegal production. The overall incentive to produce underground is reduced, so illegal firms are generally worse off after the arrival of a mafia. Thus, the size of the shadow economy decreases with the entry of a mafia. There is, however, a region where this does not hold and the government reacts by increasing public goods more than taxes. Then, legal firms actually benefit from the entry of a mafia. If the public goods are increased more than the value of a rent, then also illegal firms benefit from the entry of a mafia. The increased competition whereby the government and the mafia compete neck and neck may benefit firms! Moreover, the government’s profits exceed those of the mafia, because the government has multiple means to grab firms’ revenues.\footnote{It is straightforward to establish $G_2 > M_2$ is true under Assumption 1.} Particularly, the government is free to collect rents from firms in both sectors, while the mafia can only collect rents from firms in the illegal sector. Since the rent or the size of the shadow economy is not determined by the fine, the mafia’s payoff is unaffected by it.

Proposition 6  The arrival of a mafia increases the government’s payoff.

Proof. See Appendix B. ■

Surprisingly, the government is better off when a mafia enters the shadow economy. Indeed, as soon as the shadow economy has emerged, the government has no incentive to get rid of the mafia.

In the presence of shadow production and a mafia, government policies have the similar effects on government income as before. Taxation incurs no direct costs, but there are indirect costs through the changes in the size of the shadow economy and the tax base. From a strategic standpoint, it is wise for the government to neutralize the effect of monitoring by setting it as a base tax. What is different from the previous set-up is that rent paid to the mafia benefits the government indirectly, since the rent has the tendency to decrease the size of the shadow economy and thereby boost tax income. Thus, the government’s tax policy is supported by both its own monitoring and mafia rent-seeking.
4.2 Stackelberg game

We now make the natural assumption that a legal authority emerges before the mafia arrives. Here, we analyze a dynamic game where the government is the Stackelberg leader and the mafia is the follower. The threshold cost and the payoffs follow equations (3) - (6). Hence, optimizing for the follower yields the FOC for the mafia as in equation (13) or in terms of the best-response function

$$r = \frac{2x + 1}{4x + 1} \left( t + (b - 1) g - pz + x \right),$$  \hspace{1cm} (15)

where increases in taxation have a tendency to increase rents, while the increases in public goods or monitoring tends to decrease rents. Optimizing the government’s payoff with regard to $t, g, p$, and simplifying gives

$$\frac{\partial G}{\partial t} = \left( \frac{\partial F(\bar{c}_3)}{\partial t} + \frac{\partial F(\bar{c}_3)}{\partial r} \cdot \frac{\partial r}{\partial t} \right) (pz - t) + (1 - F(\bar{c}_3))$$

$$\frac{\partial G}{\partial g} = \left( \frac{\partial F(\bar{c}_3)}{\partial g} + \frac{\partial F(\bar{c}_3)}{\partial r} \cdot \frac{\partial r}{\partial g} \right) (pz - t) - g$$

$$\frac{\partial G}{\partial p} = \left( \frac{\partial F(\bar{c}_3)}{\partial p} + \frac{\partial F(\bar{c}_3)}{\partial r} \cdot \frac{\partial r}{\partial p} \right) (pz - t) + F(\bar{c}_3) z - p.$$  \hspace{1cm} (16)

Comparing optimality conditions in (16) to those in (12) shows that the optimal solutions for $t, g, p, r$ are different from those in Cournot. Then the solutions for unknowns are

$$t_3 = \frac{(4x + 1)(3x + 1)}{2(4x + 1) - (b - 1)^2} + z^2,$$

$$g_3 = -\frac{(b - 1)(3x + 1)}{2(4x + 1) - (b - 1)^2},$$

$$p_3 = z,$$

$$r_3 = \frac{(1 + 2x)((5x + 1) - (b - 1)^2)}{2(4x + 1) - (b - 1)^2}.\hspace{1cm} (17)$$

Evaluating these and comparing the solutions to those for the Cournot competition, we see that $t_3 > t_2, g_2 > g_3, p_2 = p_3$, and $r_3 > r_2$ (see Appendix B).

In the Stackelberg game, the government can increase the tax rate and simultaneously decrease the amount of public goods it provides. Monitoring is constant from one set-up to another as it serves as a minimal rent the government will collect from all firms. The government’s actions increase the size of the shadow economy, which forces the mafia to raise its rents since its costs are increased. Ceteris paribus,
increases in $r$ decrease $F(\bar{c})$, i.e. the costs of collecting rents are decreased. The size of the shadow economy is larger here than in the Cournot competition (see Appendix B). Strikingly, under all Assumptions

**Proposition 7** Both legal and illegal firms are worse off in the Stackelberg game than in the Cournot competition.

**Proof.** Since $g_3 - t_3 < g_2 - t_2$, legal firms must be worse off in the Stackelberg game than in the Cournot game. Since $p_2 = p_3$ and $r_3 > r_2$, illegal firms must also be worse off in the Stackelberg game. ■

Since both the legal and illegal firms lose in the Stackelberg game, it would seem that the government, the mafia, or both, benefit. Is it possible, however, for both the government and the mafia to increase their income in the Stackelberg game?

Regarding mafia income, we know that the size of the shadow economy is larger in the Stackelberg game than in the Cournot competition so rents are higher than under Cournot, implying that mafia income is higher. On the other hand, the mafia’s costs for collecting rents also go up, so the overall impact is ambiguous. For the government, we know that the amount of monitoring does not change, yet the increased size of the shadow economy income necessarily means greater income from penalties. Since the amount of public goods the government produces is lower, the costs of producing public goods decreases. Moreover, the tax rate rises, while the legal economy shrinks, so again the overall effect of the higher tax rate on the payoff is ambiguous. Comparing $G_3 > G_2$ and $M_3 > M_2$, however, shows that

**Proposition 8** Both the government and the mafia gain more income in the Stackelberg game than in the Cournot competition.

**Proof.** See Appendix B. ■

In the Stackelberg game, the government and mafia increase their payoffs. Of course, we expect here that the first-mover will capitalize on this advantage and try to grab as much as possible – behavior that results in a larger shadow economy. Since the mafia sets rents according to the size of the shadow economy, it will surely raise its rents as long as the costs for doing so do not impair the overall benefits. From the firm’s point of view, the best scenario is where the government and mafia are evenly matched to compete; if there is a leader and a follower authority, firms inevitably lose.

41
5 Welfare and enforcement

Sections 2 to 4 hold powerful implications about the government’s ability to monitor shadow production. Since the literature largely neglect the monitoring of shadow production, and hence, optimal law enforcement, let us briefly consider the welfare effects of monitoring and enforcement.

While the social welfare is a standard measure of well-being, there is no consensus on its definition with respect to individual ethics, fairness concerns nor crime. Harsanyi (1955) suggests that a public policy aimed at satisfying the preferences of the individual members of society, the welfare function should take the form of a sum of individual utilities. The problem is that, for example, a theft is just a transfer, which might not inflict any cost for the victim and as such does not affect the social welfare. In the context of shadow economy where there is no conflict of interest between the members of society and no negative externalities involved, then the illegal production must be fully accounted for in the sum of utilities and may, in fact, increase welfare.\(^{16}\) Clearly there are limits to this approach as conflicting interests in the context of criminal.

Considering the above observations, we write the ordinary social welfare defined as a sum of the utilities of firms and the authorities, i.e. the sum of costs and benefits associated with the production of goods and services in the economy. Taxes, the income from monitoring, and rents are transfers, so they cancel out. The costs of producing them, however, are apparent. As a result, we formulate the welfare effect in the presence of a mafia as

\[
W = (1 - F(\bar{c})) (1 + g) + F(\bar{c}) (1 + bg) - \int_{-x}^{\bar{c}} \frac{c}{2x} dc - \frac{1}{2} g^2 - \frac{1}{2} \beta^2 - \frac{1}{2} F(\bar{c})^2. \tag{18}
\]

We assume that the planner is now responsible for maximizing the welfare after all the players have made their decisions. If we let the planner choose the penalty, then there is only one sensible course of action:

\(^{16}\)Weiss (1976) concludes that a possibility of cheating on income taxes and random taxes in form of a probabilistic penalty might be socially useful. Davidson et al. (2007) propose that, depending on the attributes of the goods supplied on the black market, the black market may actually increase welfare. Schneider (2008) presents evidence for that working in the shadow economy is often considered as a petty crime. For in-depth discussion see Kaplow and Shavell (2002).
Proposition 9  Under a Leviathan government, the planner should set the penalty to zero.

The intuition here is that since there are only costs and no benefits associated with monitoring, the planner has no incentive to punish offenders. Shadow production is a petty and victimless crime, i.e. nobody gets hurt from it. As there are net benefits associated with illegal production, i.e. \( c_j < 0 \) for some firms, it is always welfare-increasing to have a shadow economy.

While it appears that the externalities are completely absent in our model, this is not the case. Legal producers suffer from a fiscal externality as they are the sole contributors to the production of public goods. Here, the legal firms benefit from the entry of the mafia when the mafia and the government compete equally as in a Cournot competition. Even monitoring does not alter this result. This somewhat counterintuitive result becomes evident in the following analysis.

To evaluate the effect of monitoring on the optimum, we set \( p = 0 \) and redo the analysis in Sections 2 and 4. Comparing the results with and without monitoring, simply evaluating the FOCs proves that in all these set-ups \( t > t^{p=0} \) and \( g^{p=0} = g \), i.e. the tax collected is greater with monitoring while public goods production remains unaffected. This means that, even in the presence of fiscal externality, a legal firm is better off when the government does not monitor the shadow economy, i.e. \( 1 - t^{p=0} + g^{p=0} > 1 - t + g \) for all these set-ups. The incentive to engage in shadow production, in turn, must increase, since there is no cost for getting caught. The mafia’s rent, however, is unaltered by monitoring. Also the size of shadow economy is unaffected by monitoring. Then it must be the case that both types of firms are hurt by monitoring, the government benefits from monitoring, while the mafia is unaffected by it.

Proposition 10  Monitoring strengthens the government’s position over firms, but does not affect the mafia’s position. Both legal and illegal firms are worse off under a monitoring regime.

Thus, governments that are not concerned with the welfare of their citizens should be stripped of their power to monitor illegal producers, since monitoring acts as a rent-extracting tool comparable to bribery. Correspondingly, if the planner cares for the total welfare, not to mention the position of legal firms, it should set \( z = 0 \).
6 Conclusions

This paper considered the interactions of a government, an organized crime group (mafia), and firms operating in the legal and illegal sectors of the economy. Our results highlight the role of the public policy in shaping production, since the government creates the incentives for illegal production. Mafias flourish when there is an opportunity to extract rents on production in the shadow economy.

The ability of the government to monitor illegal activities does not alter the size of the shadow economy. When the government monitors illegal production, it collects higher taxes thereby in the margin the incentives for firms are unaffected. Moreover, the government has the authority to punish offenders, so it has an extra tool for extracting rents from firms. This means that a revenue-maximizing government should always monitor the shadow economy, even though the mafia is completely unaffected by monitoring. Indeed, even in the situation where there is no collusion between the government and the mafia, monitoring benefits both of them.

When discussing the winners and losers in these different set-ups, we note that, in a mafia-free environment, legal firms benefit from shifting to the shadow economy to increase profits. In a Cournot competition, however, the appearance of a mafia on the scene changes the outcome as the profitability is likely lowered for both legal and illegal firms. When the mafia collects rents in the shadow economy, illegal firms seek to abandon the shadow economy. At the same time, the government can increase the tax rate more than it increases the production of public goods, thereby harming both legal and illegal firms. If the government, however, ends up increasing the production of public goods more than it increases the tax rate (rent), both types of firms actually benefit from the entry of a mafia. Thus, the intensified competition of authorities eliminates a part of the fiscal externality of tax evasion. In a Stackelberg game, the profits of both legal and illegal firms decrease, while the payoffs for the government and the mafia increase. Here, the appearance of a mafia is strictly harmful to firms in both the legal and illegal sector. Meanwhile, the government benefits as the mafia’s activities indirectly support its policies. If the fine is viewed as a hidden tax or a bribe, then the government should be stripped of the right to punish offenders.

The study has raised several interesting questions that suggest lines of further study, including a rich vein of potential work in the area of social welfare. Indeed, what would the equilibrium look like and the social welfare consequences be if
the government and mafia colluded? Would a more welfare-oriented government be more successful in limiting the size of the shadow economy? Finally, if illegal production is a victimless crime, is it worthwhile to reduce the incentives to go underground in the absence of harm? This result would probably not hold if the government would use all the revenue to produce public goods.
A Restrictions for interior solutions

A.1 Sub-rules for the existence

Before evaluating the existence rules, we derive several auxiliary proofs. Let $3x > (1 - b)^2$ hold. Then

**Corollary 11** It is obvious that the following statements are true $2(4x + 1) > (6x + 1) > (5x + 1) > (4x + 1) > 4x > (1 - b)^2$.

**Corollary 12** Also whenever $2x > (1 - b)$ is true then $(4x + 1) > 4x > (1 - b)$ must also be true. If $4x < (1 - b)$ then $2x < (1 - b)$.

Evaluating the limits of the penalty is useful for the later comparisons.

**Corollary 13** Comparing $z_2^2 > z_3^2$ reduces to $\frac{(b-1)+2x}{4x-(b-1)} x < \frac{(2x+(b-1))(1+3x)}{(6x+1)-(b-1)^2}$, which is true for all $2x > (1 - b)$. The contrary is true when $2x < (1 - b)$.

**Corollary 14** Comparing $z_2^3 > z_3^2$ reduces to $\frac{2x+(b-1)}{(6x+1)-(b-1)^2} x < \frac{4x+1+(b-1)}{2(4x+1)-(b-1)^2}$. Rearrange to $(2x+b-1) (2 (4x+1) - (b-1)^2) - (4x+1+(b-1)) ((6x+1) - (b-1)^2) < 0$, which is equal to $- (2x+1) (- (b-1)^2 + 4x + 2 - b) < 0$. This is true for all $x, b$.

It appears that for the existence of the interior solution in both Cournot and Stackelberg games, the penalty must be less than $z_3^2$. It is worth noticing, that when $2x < (1 - b)$ and $z_2^3 > z_3^2$, then $z_2^3 < z_3^2$ since $\frac{(b-1)+2x}{4x-(b-1)} x < 0$ while $0 < \frac{(4x+1)+(b-1))(3x+1)}{2(4x+1)-(b-1)^2} < 1$. We prove

**Proposition 15** $\frac{(4x+1)+(b-1))(3x+1)}{2(4x+1)-(b-1)^2} < 1$, i.e. $(4x+1 + (b-1))(3x+1) > 2 (4x+1) - (b-1)^2$. First, it is obvious that $|b-1| > |-(b-1)^2|$ for $0 < b < 1$. Then it follows that $4x + 1 + (b-1) < (4x+1) - (b-1)^2$. Multiply through by $(3x+1)$ and compare $(3x+1) (4x + 1) - (3x + 1) (b-1)^2$ with $2(4x+1) - (b-1)^2$. The comparison of weights of positive and negative terms yields $(3x+1) \leq 2$ while $(3x+1) \geq 1$ for all $x \in (0,1)$. Then it must be that $2 (4x+1) - (b-1)^2 > ((4x+1) + (b-1))(3x+1)$, which completes the proof.
A.2 The regions

We evaluate the restrictions necessary to have an interior solution in each game (benchmark, Cournot and Stackelberg) under three existence rules: legal and illegal production must exist, and the government must produce public goods. Although these restrictions seem to change from one set to another, there is a sub-set of the restrictions where all games have an interior solution. Inserting solutions in (9) to $F(c_1)$ gives the size of the shadow economy in benchmark, $\frac{3x-(b-1)^2}{4x-(b-1)^2}$, while inserting (14) to $F(c_2)$ gives the size of the shadow economy in the Cournot game, $\frac{3x-(b-1)^2}{(6x+1)-(b-1)^2}$, and inserting (17) to $F(c_3)$ gives the size of the shadow economy in the Stackelberg game, $\frac{(5x+1)-(b-1)^2}{2(4x+1)-(b-1)^2}$.

For all, it must be that $0 < F(c) < 1$.

In the first two set-ups for the existence of the shadow economy, it must be that $3x > (b-1)^2$, i.e. $3x > (b-1)^2$ is true. As shown in Corollary 11, the shadow economy also exists in the Stackelberg game.

Then for the existence of the public goods in the first set-up, it is required that $x > (b-1)^2$. In the second set-up, the existence of public goods requires that $-\frac{(b-1)(1+3x)}{(6x+1)-(b-1)^2} > 0$, or $(6x+1) > (1-b)^2$, while in the third set-up, it requires that $-\frac{(b-1)(3x+1)}{2(4x+1)-(b-1)^2} > 0$ or $2(4x+1) > (1-b)^2$. All of these requirements are inferior to $3x > (b-1)^2$, i.e. they are satisfied (19) is satisfied as was suggested in Corollary 11. Thus, it is sufficient to have $3x > (b-1)^2$ for interior solutions for public goods and for the shadow economy in all these set-ups. Indeed, it turns out that (19) is the sufficient condition for the existence of interior solutions of all three games.

Finally, there must be existence rules for legal production, i.e. $\pi_{ij} > 0$. Evaluating these existence rules is a bit more complicated, since they set limits on the penalty. First, we evaluate these under an assumption that $3x > (b-1)^2$. Then we write Assumption A1 for our benchmark game without the mafia, Assumption A2 for the Cournot game and Assumption A3 for the Stackelberg game. To have $\pi_{ij} > 0$ requires that

A1 Whenever $2x < (1-b)$, we have $\pi_{i1} > 0$. Whenever $2x > (1-b)$, it must be that $z^2_1 < 1 - \frac{(b-1)+2x}{4x-(b-1)^2}x$, i.e. the penalty cannot be excessive.

A2 As in the first set-up, whenever $2x < (1-b)$, we have $\pi_{i2} > 0$. When the contrary is true, we must have $z^2_2 < 1 - \frac{(2x+(b-1))(1+3x)}{(6x+1)-(b-1)^2}$.
A3 To have $\pi_{l3} > 0$, we must have $z_3^2 < 1 - \frac{((4x+1)+(b-1))(3x+1)}{2(4x+1)-(b-1)^2}$.

Comparing the existence rules for the legal sector from A1 to A3, we find that for $2x < (1-b)$ we need to set restrictions for the penalty only in the Stackelberg game. Whenever $2x > (1-b)$, it must be that one of the three upper bounds for the penalty is enforced. Corollaries 13 and 14 suggest that the most restrictive is the requirement $z_3^2 < 1 - \frac{((4x+1)+(b-1))(3x+1)}{2(4x+1)-(b-1)^2}$ to be binding.

Assuming that $3x > (b-1)^2$, we have been able to identify the rules that provide the existence of an interior solution within three regions. Therefore, we write Assumptions 1, 2a and 2b as

1. Assume $2x > (1-b)$ and $z^2 < 1 - \frac{((4x+1)+(b-1))(3x+1)}{2(4x+1)-(b-1)^2}$ are both true (Region 1).

2a. Assume $2x < (1-b)$ and $z^2 < 1 - \frac{((4x+1)+(b-1))(3x+1)}{2(4x+1)-(b-1)^2}$ are both true (Region 2a).

2b. Assume $2x < (1-b)$, $4x < (1-b)$ and $z^2 < 1 - \frac{((4x+1)+(b-1))(3x+1)}{2(4x+1)-(b-1)^2}$ are all true (Region 2b).

B Proofs

B.1 Cournot vs. No mafia

To show that the tax in a Cournot competition is higher that the tax in the benchmark, and setting $t_2 > t_1$

$$\frac{2x(1+3x)}{(6x+1)-(b-1)^2} + z^2 > \frac{-2x^2}{-4x+(b-1)^2} + z^2$$
$$\frac{(1+3x)}{(6x+1)-(b-1)^2} > \frac{x}{4x-(b-1)^2}$$
$$\frac{(1+3x)(4x-(b-1)^2)}{(2x+1)(-b-1)^2+3x} > 0,$$

is true always.
Comparing the amount of public goods as \( g_2 > g_1 \) reduces to
\[
\frac{(b-1)(1+3x)}{(6x+1)-(b-1)^2} \frac{(1+3x)}{(6x+1)-(b-1)^2} > \frac{b(1+3x)}{4x-(b-1)^2}
\]
which must be true since (20) holds. The size of the shadow economy \( F(c_1) \) is derived by plugging (9) into (4) and \( F(c_2) \) by plugging (14) into (4). Setting \( F(c_1) > F(c_2) \)
\[
\frac{3x-(b-1)^2}{4x-(b-1)^2} > \frac{3x-(b-1)^2}{(6x+1)-(b-1)^2}
\]
is true.

The profits of the legal firm are derived by substituting (9) and then (14) into (1). Setting \( \pi_{11} > \pi_{12} \)
\[
1 - z^2 + \frac{(b-1)+2x}{-4x+(b-1)^2} > 1 - z^2 - \frac{(2x+(b-1))(1+3x)}{(6x+1)-(b-1)^2}
\]
this result is subject to that of Corollary 13. It is true whenever \( 2x > (1-b) \), i.e., under Assumption 1.

For the illegal firm we show that \( \pi_{11} > \pi_{12} \), which reduces to showing that \( bg_1 > b \)
\[
\frac{x(b-1)b}{-4x+(b-1)^2} > \frac{(b-1)(1+3x)b}{(6x+1)-(b-1)^2} - \frac{(2x+1)(3x-(b-1)^2)}{(6x+1)-(b-1)^2}
\]
Evaluating
\[
((6x+1)-(b-1)^2)x(b-1)b < \left[(b-1)(1+3x)b + (2x+1)(3x-(b-1)^2)\right](4x-(b-1)^2)
\]
and simplifying yields \(-2x+1)((3x-(b-1)^2)(4x+(b-1))<0\), which is true under Assumption 1, but not true under Assumption 2 if \( 4x < 1-b \).
The payoffs for the government and the mafia are

\[
G_2 = \frac{1}{2} \left[ \frac{(3x + 1)^2 (4x - (b - 1)^2)}{[(6x + 1) - (b - 1)^2]^2} + z^2 \right],
\]

\[
M_2 = \frac{(4x - a + 2) (3x - (b - 1)^2)^2}{2 [(6x + 1) - (b - 1)^2]^2}.
\]

where \(0 < a < 2x + 1\). Comparing the government benefits \(G_2 > G_1\),

\[
\frac{1}{2} \left[ \frac{(3x + 1)^2 (4x - (b - 1)^2)}{[(6x + 1) - (b - 1)^2]^2} + z^2 \right] > \frac{1}{2} z^2 + \frac{x^2}{2 (4x - (b - 1)^2)} \left\frac{(3x + 1)^2 (4x - (b - 1)^2)}{[(6x + 1) - (b - 1)^2]^2} \right\ \frac{x^2}{(4x - (b - 1)^2)^2},
\]

is true since (20) holds.

### B.2 Cournot vs. Stackelberg

Comparing the optimal solutions in a Cournot game to those of a Stackelberg game, and proposing \(t_3 > t_2\)

\[
\frac{(4x + 1) (3x + 1)}{2 (4x + 1) - (b - 1)^2} + z^2 > \frac{2x (1 + 3x)}{(6x + 1) - (b - 1)^2} + z^2
\]

\[
\frac{(4x + 1) [(6x + 1) - (b - 1)^2]}{2 (4x + 1) - (b - 1)^2} > \frac{2x}{(6x + 1) - (b - 1)^2} [2 (4x + 1) - (b - 1)^2]
\]

\[
- (2x + 1) (b - 1)^2 > - (2x + 1) (4x + 1)
\]

is true as proposed by Corollary 12.

Comparing the amount of public goods \(g_2 > g_3\) shows that

\[
\frac{(b - 1) (1 + 3x)}{(6x + 1) - (b - 1)^2} > \frac{(b - 1) (3x + 1)}{2 (4x + 1) - (b - 1)^2}
\]

\[
\frac{1}{(6x + 1) - (b - 1)^2} > \frac{1}{2 (4x + 1) - (b - 1)^2}
\]

is true as proposed by Corollary 13.
Comparing the rents suggests that $r_3 > r_2$

\[
\frac{(1 + 2x)((5x + 1) - (b - 1)^2)}{2(4x + 1) - (b - 1)^2} > \frac{(2x + 1)(3x - (b - 1)^2)}{(6x + 1) - (b - 1)^2}
\]

is true. This is also enough to prove that $F(\bar{c}_3) > F(\bar{c}_2)$ as

\[
\frac{(5x + 1) - (b - 1)^2}{2(4x + 1) - (b - 1)^2} > \frac{3x - (b - 1)^2}{(6x + 1) - (b - 1)^2}.
\]

The profits of the legal firm are $\pi_{l3} < \pi_{l2}$ as

\[
1 - \frac{(4x + 1 + (b - 1))(3x + 1)}{2(4x + 1) - (b - 1)^2} < 1 - \frac{(2x + (b - 1))(1 + 3x)}{(6x + 1) - (b - 1)^2}
\]

then from Corollary 14 we know that this is true in all regimes.

For the illegal firm we show that $\pi_{l2} > \pi_{l3}$, which reduces to showing that $bg_2 - r_2 > bg_3 - r_3$. Now since $g_2 > g_3$ while $r_3 > r_2$ and $z_2 = z_3$ then this is always true.

The payoffs for the government and the mafia are

\[
G_3 = \frac{(3x + 1)^2}{2(2(4x + 1) - (b - 1)^2)} + \frac{1}{2} z^2,
\]

\[
M_3 = \frac{(4x - a + 2)((5x + 1) - (b - 1)^2)}{2(2(4x + 1) - (b - 1)^2)^2}.
\]

To show that $G_3 > G_2$

\[
\frac{(3x + 1)^2}{2(2(4x + 1) - (b - 1)^2)} + \frac{1}{2} z^2 > \frac{(3x + 1)^2 (4x - (b - 1)^2)}{2((6x + 1) - (b - 1)^2)^2} + \frac{1}{2} z^2
\]

then

\[
\frac{1}{(2(4x + 1) - (b - 1)^2)(4x - (b - 1)^2)} > \frac{1}{((6x + 1) - (b - 1)^2)^2}
\]
reduces to 

\[ ((6x + 1) - (b - 1)^2)^2 > (2(4x + 1) - (b - 1)^2)(4x - (b - 1)^2) \]. Simplifying yields \((2x + 1)^2 > 0\) which is true for all \(x\).

To show that \(M_3 > M_2\)

\[
\frac{(4x - a + 2)((5x + 1) - (b - 1)^2)^2}{2(2(4x + 1) - (b - 1)^2)^2} > \frac{(4x - a + 2)(3x - (b - 1)^2)^2}{2[(6x + 1) - (b - 1)^2]^2}
\]

\[
\frac{((5x + 1) - (b - 1)^2)^2}{(2(4x + 1) - (b - 1)^2)^2} > \frac{(3x - (b - 1)^2)^2}{((6x + 1) - (b - 1)^2)^2}
\]

is true since \((21)\) is true.
References


Do The Catholic and Protestant Countries Differ in Their Tax Morale? *

November 12, 2009

Contents

1 Introduction 57

2 Model 61
  2.1 Tax Morale ................................. 61
     2.1.1 Laffer-curve and Tax Morale .......... 63
  2.2 Multiple Equilibria and Prisoners’ Dilemma ............................. 64
  2.3 Towards Empirical Hypotheses ........................................... 65

3 Empirical Analysis 67
  3.1 Method ...................................... 67
  3.2 Hypotheses and Data .................................................... 70
     3.2.1 Variables ................................................. 70
     3.2.2 Statistical properties ................................. 72
  3.3 Estimation Results .................................................... 74
  3.4 Caveats ............................................ 78

4 Final Remarks 78

*The paper is co-authored by Professor Vesa Kanninen. A version of this paper is forthcoming in Empirica.
1 Introduction

One of the most influential analyses of the links between the economy, religion and morality has been Max Weber’s *Protestant Ethic and the Spirit of Capitalism* from 1904. The popular view - though an imprecise one - is that it has been protestant ethics which has been benign to the work ethic, having promoted the success of capitalism. Such a proposition may also be translated into an empirical prediction of how people adhering to different religions judge illegal economic activities. For example, Schneider (2005) suggests that the shadow markets are larger in the Mediterranean countries than, say, in the northern welfare states.

In their ground breaking paper, Reckers et al. (1994) studied the interaction of tax ethics with the tax rates. An increase in the marginal tax rate seems to increase the noncompliance among those individuals who strongly disagreed with the statement that tax evasion is morally wrong. Conversely, those who condemned tax evasion as morally wrong did not alter their responses. Other researchers have, however, presented more skeptical views and pointed to group norms: if others cheat then one may cheat without feeling guilty (for discussion see Pilkington 2007). Alm and Torgler (2006) propose that there are cultural differences regarding tax morale across countries and in particular that northern Europeans have higher tax morale than southern Europeans. Studying Austria, Torgler and Schneider (2005) find that from 1990 to 1999 the tax morale has decreased, yet is high compared with other European countries. More recently, Feld et al. (2008) evaluate the tax morale in Germany after German unification. They find that the tax morale in the East
and the West Germany converged from 1990 to 1999 being almost identical in the end. Most interestingly, it was the initially higher East German tax morale that converged to the lower West German levels, i.e., the standards decreased.

In this paper, we ask whether a religious denomination can explain differences in tax morale and tax compliance in particular. There is indeed evidence to support this view, as some empirical studies evaluate the economic effects of religion (cf. Iannaccone, 1998). Substantial differences exist between catholicism and protestantism in both the structure of their beliefs and their enforcement mechanisms. Kirchgässner (1999) argues that historically in north Europe, state and religious authority were largely held by one person and offences against the state were therefore religious offences (and consequently a sin). Arrunada (2004) also argues that the catholic theology and practice facilitate personal transactions while protestantism favors values and types of moral and legal enforcement better adapted to impersonal trade. Guiso et al. (2003) find a relation between religion and people’s attitudes toward legal rules, such as trust in the legal system, cheating on taxes, or paying bribes. In particular, they find that the willingness to cheat on taxes is weakest in judaism followed by protestantism, catholicism, hinduism and islam.

The theory of tax evasion goes back at least to Allingham and Sandmo (1972), and the subsequent literature has linked social norms to tax evasion (see Gordon (1989)) and to punishment. Brekke et al. (2003) propose a model where a consumer’s self-image can only be improved by striving towards what she truly believes to be morally right. The experiment by Myles and Naylor (1996) supports a tax evasion model where utility is derived from a social custom when taxes are paid honestly. Fehr and Gächter (2002) indicate that people are willing to invest in public goods as long as they have the opportunity to inflict punishment on those who free ride on cooperation. Orvaska and Hudson’s (2002) results suggest that evasion is condoned by a large proportion of the population although people appear to be deterred from tax evasion by the consequences of being caught.

Recently, Katungi et al. (2006) surveyed the motives for informal work in the UK. While the single most significant reason for informal work was to alleviate poverty the respondents gave also other reasons: employers are avoiding paperwork in times of critical staff shortages (in the construction industry and in the clean-
ing services in particular), respondents took informal second job to top up their low wages, foreign students and asylum seekers make their living in informal employment, graduates perceived lack of formal jobs matching their skills, and some respondents are avoiding the application costs of formal jobs. In general, many respondents saw the informal job as a route to the formal one.

Our paper contributes to the literature on tax morale and shadow markets. We introduce a model of consumers and entrepreneurs who value their moral code but are inherently opportunists in a sense that their morality is priced in market transactions. In our model, entrepreneurs provide their services to consumers either through the legal or illegal market. Tax morale is reflected in the difference in the private valuation of the legal and illegal transactions, arising from the lack of self-respect consequent on tax evasion.¹ A legal producer benefits from the moral aspirations of consumers', which provide a compensatory mechanism for the tax to be paid. An illegal producer avoids the tax and can contract with the consumers visiting the shadow market at a lower price. It is not that in equilibrium all potential entrepreneurs are active. They have an outside option in terms of benefits when choosing unemployment. Tax evasion results in a fiscal externality in the financing of public goods, which have the property that they enhance the private productivity of entrepreneurs. The tax rate chosen by a revenue-maximizing government hinges upon the tax morale of people in the economy.

While there exist both direct and indirect approaches for quantifying the size of the shadow economy, structural equation models which incorporate latent variables have become a standard technique for estimating the size of the shadow economy.² The direct approaches make use of microeconomic data from tax auditing or surveys, while indirect approaches typically dwell on discrepancies on the macroeconomic measurements such as national expenditure and income statistics or official and actual labor force. On the use of the former approach there are two potential problems. Firstly, the tax auditing data is difficult to come by and the auditions are often concentrated to one problematic sector, whereby the results are not general. Second, in surveys the respondents tend to lie in such delicate matters, therefore

¹Hausman and McPherson (1993) have provided a review of why and how morality influences economic outcomes. Cf. Frank (1987) for a pioneering analysis of honesty and dishonesty.
²For a review see Schneider and Enste (2000) and Lyssiotou et al. (2004).
even if they were available the results are not necessarily very reliable. On the use of the latter approach the problem has been that the studies typically rely on one source of information, i.e. they concentrate on the discrepancy between two indicators instead of many. The structural equation approach overcomes this fault, although the existing studies have been evaluated critically by Breusch (2005 a,b) among other. Three of his points are more critical, and while we are not able to correct all the flaws he demonstrates the previous studies have suffered, we take several attempts to control for them. Also their impacts on our results are discussed.

We work with the data from the OECD to evaluate the effects of tax, government consumption, social security, and interest rates on the shadow economy. In our data, we identify two regimes. In the first regime, 1979-1992, the activities of public sector grew than during the latter regime, 1992-2003. To test the hypothesis that religion affects tax morale and hence tax evasion, we use the data from the World Values Surveys to group the countries according to their primary religious denomination. The countries form natural groups of the catholic south and the protestant north, whereby the other countries make up the reference group that facilitates the comparison between groups and with the previous studies. Should this hypothesis be valid, it will be reflected in the elasticity of the size of the shadow economy with regard to tax.

We report three major findings. First, there are some differences in tax morale across religions, but for the latter regime we reject the hypothesis that countries with a different religious background exhibit different tax morale. Our estimates suggest that within the OECD, the norms are global. Second, even though we do not calibrate the index of the size of the shadow economy, our approach with direct effects included suggests that the size of the shadow markets, however, has been overestimated by previous studies. Third, our results support the view that there was a regime switch in the evolution of the public sector in the early 1990s and that the underlying model of the shadow economies within these regimes differs.

The road map of our paper is that we first develop our theoretical model in Section 2. The model is estimated in Section 3, which also reports the econometric results. Section 4 concludes.
2 Model

2.1 Tax Morale

Does the tax morale differ across countries? It is possible that private incentives to avoid taxes are sufficiently strong promoting shadow activities. We show first that countries can diverge to different levels of shadow activities, or they can converge to the same level of shadow activities. Of course, also differences in policy can cause differences in the shadow activities. Then we study the conditions as to when it is the prisoners’ dilemma which in the first place creates incentives for shadow activities.

Our model is a version of the seminal vertical product differentiation model introduced by Mussa and Rosen (1978). To fix the ideas, we consider an economy where consumer services can be bought either in a legal or an illegal market. The mass of consumers is normalized to unity. The number of producers is normalized to $\eta$, where $\eta > 0$. Each consumer buys one service from either one of the markets at most while each producer can service several buyers. Legal producers are subject to a tax, $\tau > 0$. Illegal producers can contract with the buyers at a lower price and abstain from paying the tax.

People are assumed to be heterogeneous in terms of their moral aspirations, reflected in their willingness to pay. The consumer with the highest moral standard values the transactions with a legal producer by $\alpha_1$ and the transactions with an illegal producer by $\alpha_2$ where $\alpha_1 > \alpha_2 > 0$. This assumption captures the idea that a deviation from the social norm results in stigma or a loss of self-respect. The consumers are ordered with a linearly declining valuation of the goods.

Thus, if the consumer $i$ buys from the legal and consumer $j$ from the illegal producer, their net (indirect) utilities, reflected in their willingness to pay are\(^3\)

\[
\begin{align*}
    u^i &= \alpha_1 (1 - i) - p_1; \\
    v^j &= \alpha_2 (1 - j) - p_2; \\
    \alpha_1 > \alpha_2 > 0,
\end{align*}
\]  

where $p_1, p_2$ are the market prices of the legal and illegal products, yet to be determined.

\(^3\)For introduction of such utilities, cf. Katz and Shapiro (1985).
For the sake of simplicity, the entrepreneurs have a choice between operating in the legal or in the illegal market. Those choosing to operate in the illegal market face a moral cost. Therefore we introduce the utility functions of the entrepreneurs as

\[ U^l = \pi_1; \quad V^l = \beta \pi_2, \]  

where the profits they earn are denoted by \( \pi_1, \pi_2 \) and where \( \beta < 1 \) is the valuation of the illegal profit.\(^4\) The legal and illegal outputs will be denoted by \( x_1 \) and \( x_2 \). Public goods denoted by \( g \) will enhance private productivity. Inactive entrepreneurs are entitled to a government transfer in terms of unemployment benefit, \( \delta > 0 \), representing an opportunity cost. We introduce linear revenue functions and quadratic cost functions. Hence, the marginal costs are increasing in output. Then the profits of the legal and illegal entrepreneurs are

\[ \pi_1 = p_1 g x_1 - \frac{1}{2} x_1^2 - \tau; \quad \pi_2 = p_2 g x_2 - \frac{1}{2} x_2^2. \]

Product markets are competitive and the entrepreneurs are price-takers. There is free entry into both markets. The entry conditions are given by \( \pi_1 \geq \delta, \beta \pi_2 \geq \delta \) where the unemployment compensation is used to measure the opportunity cost.\(^5\) The first-order conditions are \( x_1 = gp_1 \) and \( x_2 = gp_2 \). Market prices adjust to eliminate the rents. From \( \pi_1 = \beta \pi_2 = \delta \), we obtain the prices and the first-order conditions give the optimal outputs

\[ p_1 = \frac{\sqrt{2(\tau + \delta)}}{g}; \quad p_2 = \frac{\sqrt{2\delta/\beta}}{g}; \quad x_1 = \sqrt{2(\tau + \delta)}; \quad x_2 = \sqrt{2\delta/\beta}. \]

Provided that the tax rate exceeds the transfer by a margin, \( \tau > \delta (1 - \beta) / \beta \), the illegal price and output are lower than the legal price and output. The private benefits from tax evasion are split between the buyers and the legal entrepreneurs.

\(^4\)Note that there is no cost of being caught operating in shadow economy in the model. Introducing such a cost would affect the equilibrium and reduce illegal output.

\(^5\)It is assumed that the illegal producers are not able to exploit government transfers if they are active. Introducing the exploitation of government transfers by the illegal producers would be trivial.
Moreover, in equilibrium, those entrepreneurs operating in the illegal sector have a greater profit than those operating in the legal sector, $\pi_1 = \delta < \pi_2 = \delta/\beta$. Yet, their utilities are equalized as the illegal producer suffers a moral cost.

To solve for the aggregate legal and illegal production, we denote by $m$ the marginal consumer who is indifferent between buying the legal and the illegal product. Similarly, we denote by $n$ the consumer who is indifferent between buying the illegal product and none. The consumers in the interval $(0, m)$ buy the legal product and consumers in the interval $(m, n)$ buy the illegal product. The net utility of the consumer $n$ is zero, i.e. $\alpha_2 (1 - n) - p_2 = 0$. Then the total number of customers, including those visiting the legal and those visiting the illegal markets is $n = 1 - p_2/\alpha_2$.

It holds for the marginal consumer $m$ that $\alpha_1 (1 - m) - p_1 = \alpha_2 (1 - m) - p_2$. This can be used to solve for the number of customers in the legal sector,

$$m = 1 - \frac{p_1 - p_2}{\alpha_1 - \alpha_2} = 1 - \frac{\sqrt{2(\tau + \delta)} - \sqrt{2\delta/\beta}}{g(\alpha_1 - \alpha_2)}.$$ 

Then, the size of the shadow sector, measured by the number of its consumers is

$$n - m = \frac{p_1 - p_2}{\alpha_1 - \alpha_2} - \frac{p_2}{\alpha_2} = \frac{\sqrt{2(\tau + \delta)} - \sqrt{2\delta/\beta}}{g(\alpha_1 - \alpha_2)} - \frac{\sqrt{2\delta/\beta}}{g\alpha_2}. \quad (4)$$

From the condition $n - m > 0$, we obtain

**Lemma 1** The condition for the existence of the shadow sector is $\alpha_2/\alpha_1 > \sqrt{\delta/\beta(\tau + \delta)}$.

Moreover, the weaker the tax morale of consumers (small $\alpha_1 - \alpha_2$) or producers (great $\beta$) is or the greater the tax rate $\tau$ is, the greater is the shadow economy.

The condition for the legal sector to survive is $m > 0$, or $\alpha_1 - \alpha_2 > \frac{1}{g} \left( \sqrt{2(\tau + \delta)} - \sqrt{2\delta/\beta} \right)$. Thus, only sufficiently strong tax morale can sustain the survival of the legal markets. In terms of a ceiling to the tax rate and transfer payments (unemployment compensation), this amounts to $\tau^{\max} < \frac{1}{2} \left( g(\alpha_1 - \alpha_2) + \sqrt{2\delta/\beta} \right)^2 - \delta$. Clearly, a too high tax will kill the legal production.

### 2.1.1 Laffer-curve and Tax Morale

We now follow the tradition in public finance which considers a government as a revenue-maximizing Leviathan, extracting a share $\lambda > 0$ for its own consumption.
It thus maximizes $\lambda T$, where $T = \tau m/x_1 = g + u\delta$ is the tax revenue available for the government to produce public goods and provide benefits. In Appendix A, we show the following result:

**Proposition 2** The economy faces a concave Laffer curve with the revenue maximizing tax rate satisfying $\tau^* < \tau^{\text{max}}$.

A small valuation difference $\alpha_1 - \alpha_2$ points to low tax morale whilst the large difference points to high tax morale. An increase in tax morale enhances the government’s ability to collect tax revenue, since $\partial T/\partial (\alpha_1 - \alpha_2) > 0$ and $\partial \tau^{\text{max}}/\partial (\alpha_1 - \alpha_2) > 0$ for any given $g$. This means that with increased tax morale, the Laffer curve moves up and to the right.

The policy variables are inter-related through the government budget constraint. An increased tax rate raises the equilibrium price in the legal market, $\partial p_1/\partial \tau > 0$, making each legal producer bigger, $\partial x_1/\partial \tau > 0$. However, the number of legal producers declines, $\partial (m/x_1)/\partial \tau < 0$. As a result, there will be more unemployment. Now, the number of illegal producers, $(n - m)/x_2$ will increase, as $\partial (n - m)/\partial p_1 > 0$, reducing unemployment. Which of two effects on unemployment dominates? Evaluating the total effect and substituting the optimal production solutions, we find that $\partial u/\partial \tau < 0$ when $\tau > \delta \left(\frac{1-\beta}{\beta}\right)$ i.e., when the output and price are lower in the shadow sector than in the legal sector. Suppose that the economy is settled at the top of the Laffer curve with the property that $\partial T/\partial \tau = 0$. Thus, an increase in the tax rate cannot generate additional tax revenue. However, an increase in the tax rate increases the number of those eligible for unemployment compensation.\(^6\)

### 2.2 Multiple Equilibria and Prisoners’ Dilemma

How does a shadow economy arise in the first place? Consider an economy in an initial equilibrium with no illegal sector all entrepreneurs having a profit $\pi_1 = \delta$. All producers pay the tax. It now holds for the marginal consumer that $\alpha_1(1-m)p_1 = 0$. Therefore, the number of buyers is $m = 1 - p_1/\alpha_1$. The output and the price of

\(^6\)Complications may arise if the economy is not settled down at the top of the Laffer curve. Second, with access to public debt, the government may actually face a soft budget constraint.
the legal firm are the same as above. In the absence of the illegal market, there are more legal producers. This means that the government’s tax revenue is greater, too.

The amount of unemployment compensation and the public goods to be supplied in the equilibrium are

\[ \delta u = \delta \left( \eta - \frac{m}{x_1} \right), \quad g = T - \delta \left( \eta - \frac{m}{x_1} \right). \]

If each entrepreneur expects everyone else to commit to contributing to the tax revenue and financing of public goods, the economy can settle down in a "good equilibrium". If each entrepreneur, however, expects that the others will not contribute to the tax revenue, the economy can settle down in a "bad equilibrium" with the substantially reduced supply of public goods. Though no one desires the economy to settle down in the bad equilibrium every entrepreneur, however, has a private incentive to deviate as he can increase his profit by establishing shadow activities. The economy faces a prisoner’s dilemma. Thus,

**Proposition 3** Shadow economy arises as prisoner’s dilemma once the moral code - in terms of sufficiently high $\beta$ - is violated.

**Proof.** A legal producer generates a profit $\pi_1 = (gp_1)^2 / 2 - \tau$. Any legal producer can, however, make a bigger profit by abstaining from the tax payment, provided no other producer follows and the consumers cannot monitor the tax evasion. Such an incentive arises for each producer whose valuation of the legal production satisfies $\pi_1 = (gp_1)^2 / 2 - \tau < \beta \pi_2 = \beta (gp_1)^2 / 2$. The parameter $\beta$ has to be sufficiently high for this condition to hold. However and given this condition, as every producer has this incentive, the economy will settle down in the bad equilibrium. ■

The good equilibrium is therefore subject to private opportunism. When everyone behaves according to this incentive, the resulting reduction in the tax revenue leads to a reduction in the public goods and unemployment benefit. Thereby the utility of each producer is reduced.

### 2.3 Towards Empirical Hypotheses

What predictions are to be derived for an empirical study? First, assume that people in a country A have developed a stronger attitude towards social norms
than in country B. In terms of our model, consumers and entrepreneurs are then equipped with the following preferences

\[ \alpha_1^A > \alpha_1^B, \quad \alpha_2^A < \alpha_2^B, \quad \beta^A < \beta^B. \]

In our model, we can trace the effects of government policies on the structure of the production sector. Accordingly, we formulate the following hypotheses, to be briefly repeated in Section 3:

1. An increase in the tax rate increases the size of the shadow market, \( \partial (n - m) / \partial \tau = \frac{1}{2 \sqrt{2(\tau + \delta)}} \frac{\sqrt{2}}{\delta (\alpha_1 - \alpha_2)} > 0. \) This effect is magnified by low tax morale whereby the valuation difference \( \alpha_1^A - \alpha_2^A < \alpha_1^B - \alpha_2^B. \)

2. An increase in public goods decreases the size of the shadow sector, \( \partial (n - m) / \partial g = \frac{1}{g^2} \left( \alpha_1 \sqrt{2(\tau + \delta)} - \alpha_2 \sqrt{2(\tau + \delta)} \right) < 0, \) given the existence condition in Lemma 1. Again the tax morale magnifies this effect.

3. The unemployment benefit increases both prices, having an ambiguous effect on the size of the shadow market, \( \partial (n - m) / \partial \delta \leq 0. \) The effect is positive when \( \alpha_2 / \alpha_1 > \frac{1}{\beta} \sqrt{\beta (\tau + \delta)} / \delta \) and negative otherwise.

Since all of these effects are magnified by the low or high tax morale, we formulate the following hypothesis

4. The larger the valuation difference between the legal and illegal product, i.e. the stronger the tax morale, the more likely it is that an increase in the tax variable, a decrease in the public goods variable, and a decrease in the unemployment benefits all increase the size of the shadow sector.

For the reasons discussed in this section, various fiscal variables are included among the explanatory variables to study whether we can identify cross-country differences between the effects of the policy variables when countries have a different religious denomination. The comparative static analysis above demonstrates that the effects of the policy variables are magnified by the tax morale. The empirical section estimates these relations taking that the tax morale operates as the
transmission mechanism between those policies and shadow markets. The potential cross-country differences in tax morale are expected to show up in the estimated coefficients of tax, benefits and government consumption variables.

3 Empirical Analysis

3.1 Method

A latent variable is a random variable whose realizations are hidden from an observer. Even though it may have operational implications for relationships among observable variables. Also the shadow activities are unobservable which represents a challenge for the econometric testing of the theoretical hypotheses. The problem, however, is quite common in economics\(^7\) and can be coped by SEM (structural equation modeling). Its origins in econometrics are in simultaneous equations models and dates back to Haavelmo (1943). The identification and estimation have been discussed by Konijn (1962)\(^8\), and the estimation theory was further developed by Zellner (1970) and Jöreskog and Goldberger (1975)\(^9\).

A structural equation model consists of two parts, the first being the structural model and the second being the measurement model (Figure 1). The structural model links the observed causes to the unobserved variable. The measurement model, also known as a common factor model, links the unobserved latent variable to observed indicators. The Multiple-Indicator-Multiple-Cause (MIMIC) model is a version of SEM with one latent variable. This method has been used to estimate the magnitude of the shadow economy by Frey and Weck-Hanneman (1984), Aigner, Schneider and Ghosh (1988), Giles and Tedds (2002), and Dell’Anno and Schneider (2005) among others. For example, Schneider (2005) finds that the estimated unweighted average of the size of the shadow economy in OECD in 2002/2003 was 16% of that of official GDP, the shadow economy being the smallest in the US (8

\(^7\)The "causes" may be viewed as the instruments in IV-approach see Goldberger (1972) and Angrist et al. (1996).

\(^8\)See also Zellner and Theil (1962), Rothenberg and Leenders (1964), and Sargan (1964).

\(^9\)The special issue of the Journal of Econometrics (1983) is devoted to this approach. See also Aigner et al. (1984), Aigner and Goldberger (1979), and Schumacker and Lomax (2004).
The studies using the MIMIC model to estimate the shadow economy have been criticized by Hill (2002), Smith (2002), and Breusch (2005a,b). For example, both Hill (2002) and Smith (2002) state that none of the studies has a theory behind the estimated model. Breusch (2005a,b) points out several failures in the previous attempts to estimate the size of the shadow economy as he replicates the empirical analysis of three papers. He documents several failures starting from undocumented transformations of the data to the failure to recognize the critical statistical implications that are present in the method. The first flaw the keen advocates of the MIMIC-method fail to encounter is that the Maximum Likelihood method assumes that the variables in the model are multivariate normal. Knowing the data that is typically used in these applications this is unlikely. Thereby, the inferences from these applications might be fallible. The second is that since the reduced form model is estimated the cause variables should be (weakly) exogenous, otherwise the estimates are biased. The third flaw relates to the indicator variables and here Breusch pinpoints two potential dangers. Accordingly, a MIMIC model requires that the indicators are conditionally independent of the causes, given the latent variable. Moreover, the indicators should be mutually independent, given the latent variable. In other words, all the connections that the indicator variables have with the causal variables, and with each other, are carried through the latent variable (Breusch 2005b). These two assumptions run counter to the conventional economic theory.
To summarize, the latent variable approach is a valid method to overcome the problem of non-observables. Yet its use is not a panacea, as Breusch (2005b) demonstrates. This criticism is a valid one but it can be handled.

Suppose throughout that all variables are measured around their respective population means. Thus, the intercept is omitted. The structural model assumes that the latent variable $y^*$ is determined as

$$y^* = \alpha_1 q_1 + \alpha_2 q_2 + \ldots + \alpha_k q_k + \epsilon,$$

where the observable exogenous causes $q_1, \ldots, q_k$ are obtained from the theoretical model and where $\epsilon$ is the disturbance term.

In our measurement model, the latent variable determines, together with (some) cause variables and disturbances $u_1, \ldots, u_m$, the set of observable indicators $y_1, \ldots, y_m$

$$y_1 = \beta_1 y^* + \Sigma_{i=1}^k \gamma_{1i} q_i + u_1, \ldots, y_m = \beta_m y^* + \Sigma_{i=1}^k \gamma_{mi} q_i + u_m.$$

The earlier studies imposed the constraint $\gamma_{ji} = 0$ for all $j, i$. We, however, estimate the specification which allows for some non-zero $\gamma$’s, i.e. allows for direct effects.

The indicator variables are assumed to reflect the evolution both of the unobserved latent variable and the cause variables and are thought to covariate with them. In vector-form

$$y^* = \alpha' q + \epsilon,$$

$$y = \beta y^* + \gamma q + u,$$

with $E[\epsilon] = 0, E[u] = [0 0 \ldots 0]'$, $E[\epsilon u] = [0 0 \ldots 0]'$, $E(\epsilon^2) = \sigma^2$ and $E(u \times u') = \Theta^2$.

Substituting, we obtain the reduced-form relation

$$y = (\beta \alpha' + \gamma) q + \beta \epsilon + u = \Pi' q + v,$$

where the reduced-form coefficient matrix is $\Pi = \alpha \beta' + \gamma'$ and where the reduced-form disturbance vector, $v = \beta \epsilon + u$ has covariance matrix $\Omega = E(vv') = \sigma^2 \beta \beta' + \Theta^2$.

We estimate the coefficient matrix $\Pi$ in equation (9) which contains only observable variables. Estimation of structural parameters is computed using full information maximum likelihood technique\textsuperscript{10}, making use of the restrictions implied

\textsuperscript{10}For the identification of simultaneous equation systems and FIML estimation see Intriligator et al. (1996).
in both the coefficient matrix $\Pi$ and the covariance matrix of the error term $\nu$. Such restrictions are needed because without them the model would be under-identified. To determine $\alpha$ and $\beta$, one of the factor loadings has to be fixed and it has become a common practice to set $\beta_1 = 1$.

With all direct effects included, even those restrictions would not suffice to identify the model. The identification of direct and indirect effects has been discussed by Fox (1980) and Bollen (1987)\textsuperscript{11}. The \textit{a priori} restrictions for direct effects, i.e. restrictions that some of the $\gamma$’s are zero while others are not, are based on economic theory and to our model.

### 3.2 Hypotheses and Data

#### 3.2.1 Variables

There are differences among countries according to their historical, regional and cultural heritage, which might be reflected in the evolution of the shadow economy. Combining the evidence from the four waves of the World Values Surveys (1981-2005), we have learned that those belonging to a religious denomination are on average more likely to disapprove such behavior as cheating on taxes, claiming government benefits when not entitled to, and accepting bribes. In particular, protestants are more likely adhere to their strict values than catholics.

Subsequently, we group countries according to their dominating religion, where two indicators of the religious orientation are taken into account. The first one is the proportion of the citizens of a country who belong to a religious denomination, while the second is the majority religious denomination of a country. Using these two measures, there are two distinctive groups: the catholic south (CS) with a membership rate of 80 - 95 per cent of the population (Greece, Italy, Spain and Portugal) and the protestant north (PN) with a membership rate of 80 - 90 per cent (Denmark, Finland, Norway and Sweden). The remaining OECD countries form a control group, the rest of the world (RoW; Australia, Austria, Belgium, Canada, Germany, France, Ireland, Japan, Netherlands, Norway, New Zealand, Switzerland, UK, the US). From the previous studies and WVS we conclude that

\textsuperscript{11}See also Bielby and Hauser (1977) for the review of the issues.
protestant north should obey the moral code of not cheat more strictly than the catholic south. The control group should give us a reference point on where the remaining, culturally more heterogenous countries are.

For the empirical analysis, we identify four cause variables and three indicators (see Appendix B). The hypotheses will be tested using annual data for 21 industrialized OECD countries for 1979 through 2003. All variables, except the nominal interest rate, are measured as growth rates per annum.

**Cause variables**

1. Total tax revenue \((\tau)\). The higher the tax burden, the greater the shadow economy. It thus reflects both the tax evasion effect and the strength of tax morale.

2. Government consumption \((g)\) as a measure of the public goods. According to our model, the growth of the public good’s production is expected to decrease the growth of the shadow economy.

3. Government transfers \((\delta)\). We expect the transfer payments to reduce the incentive to work in the shadow economy.

4. Nominal short-term interest rate \((r)\). An increase in the interest rate increases the opportunity costs of holding money, thus reducing shadow economy.

**Indicators**

1. Real GDP per capita \((x)\). From our theoretical model we expect that higher demand will increase the size of both sectors, i.e., when the legal economy is booming, the unofficial sector is also booming.

2. Labor force participation rate in the official sector \((l)\). There is evidence that in industrialized countries, those who moonlight are already a part of the labor force, contrary to the case of developing economies\(^{12}\). Therefore, we expect

\(^{12}\)Katungi et al. (2006) surveyd the motives for informal work in the UK. The single most important reason for informal job was to alleviate poverty, while many respondents saw the informal job as a route to the formal one.
that increases in moonlighting and shadow activities will increase the labor force by encouraging people to register for the official labor market.

3. Real currency in circulation per capita \((m)\). This variable measures changes in payment habits. The greater growth of cash payments reflects the greater growth of the shadow transactions.

### 3.2.2 Statistical properties

Based on the Levin-Lin-Chu (2002) and Im-Pesaran-Shin (2003) tests, we conclude that these series are stationary. Since the series are non-stationary at their levels, we prefer to use the differences to avoid spurious regression. Moreover, scrutinizing the public sector variables suggests that there are breaks in the data (see Appendix B for details). Hence, we split the data into two periods, calling the first the regime of public sector growth (1979-1992) and the second regime of a mature public sector (1992-2003).

ML-estimation infers that the data should be normal, which considering the data in hand might not be feasible. At Table 1, we report the skewness, kurtosis and Jarque-Bera test statistics for each of our variables. First of all, with the exception of real GDP per capita, all data seem to be right tailed since the skewness is non-zero. Second, all data seems to be leptokurtic, or peaked, since for all variables the kurtosis is above three. Therefore, Jarque-Bera test is statistically significant for all variables, markedly so for real currency per capita. Hence, the data is not univariate normal, and likely not multivariate normal either which, if not accounted for, will affect the estimated standard errors and the related test statistics. Since the methods that rely on the normality of the data typically underestimate the true standard errors in the presence of positive kurtosis, the hypothesis testing based on those may be misleading.

There are methods that correct standard errors in the presence of non-normality, for instance bootstrapping although it is not a panacea for small samples. In small samples, the standard error estimates are almost always unreliable, yet also size corrections may produce biased estimates for standard errors (for details see Schermelleh-Engel et al., 1998). Foremost, the current bootstrapping procedures cannot handle missing data, which occurs in our first period. As the method is not

72
admissible in the presence of missing data, we resort to full information maximum likelihood, knowing that hypothesis testing might be flawed.

We address the endogeneity of the variables with Granger-causality test and by estimating a simple general model to control for the impact of multiple causes (for details see Appendix C). According to Granger causality tests, there is some evidence that the proposed causes and indicators are not exogenous. In case of causes the inference from the general model, which includes lags of other cause variables and in particular two lags for each indicator variable. For instance, it appears that real GDP per capita significantly affects all our cause variables, while some of the causes are also affected by another indicator. The causes are not strictly exogenous, which will have an impact to our results. The more interesting result is that, controlling for the cause variables as instruments for the shadow economy, our indicators seem to be independent on each other. Only the labor force participation seems to be affected by the first lag of real GDP per capita. Since the bias introduced by the fact that the indicator variables are not exogenous is far greater than the bias due to endogeneity of the cause variables, this is a good sign.
3.3 Estimation Results

To identify the model, we impose two kinds of restrictions as discussed above. First, to fix one of the factor loadings, we impose the restriction $\beta_1 = 1$, suggesting that an increase in the shadow economy increases real GDP per capita, both in growth terms. Second, we set some of the direct effects to zero, seeking the parsimonious representation of the model. Since the model is not identified if too many direct effects are included at the same time, we seek the ones that are significant by an iterative process guided by the $t$-statistic and the fit criteria. We first add the direct effects one by one, keeping those that are significant. We iterate the process starting from alternative direct effects. Note that this type of iteration probably distorts the nominal significance level of the model.

The method does not allow for testing whether the same structure applies from one period to another in a multiple group setting. Hence, we estimate the models for the two periods separately. As the models are, however, nested in a particular period, we may test whether $\beta$’s and $\alpha$’s differ across groups of countries. In the first regime, the default model allowing for the differences in $\beta$’s is accepted, indicating that the effects of the shadow economy on our indicator variables differ across groups of countries. In the second regime, we reject the default model; i.e., $\beta$’s are equal across country groups. We also reject the hypothesis that $\alpha$’s are equal across the groups.

The estimation procedure converges, yielding a positive semidefinite covariance matrix. The diagnostic test results for the two regimes are mixed, yet provide some support for the model(s). For the first regime, the value of the $\chi^2$-test with 33 degrees of freedom is 33.32 ($p$-value 0.00), which is non-supportive for the model, whereas for the latter regime it is 28.67 with 28 degrees of freedom ($p$-value is 0.43) which is supportive. Since the $\chi^2$-test is sensitive to the sample size and the non-normality of the variables, we resort to other fit criteria, where TLI in particular has proven to be least affected by the biasing factors (see Appendix D). The fit indices for the first regime are as follows: NFI 0.96, RFI 0.77, TLI 0.86 and CFI 0.98, while the root mean square error of approximation is 0.07. The indices are

---

13 Since we minimize the distance between the sample and model implied covariance matrices, it is important that the minimum is achieved.
slightly better in the latter period, as NFI 0.93, RFI 0.84, TLI 0.99 and CFI 0.99, while the root mean square error of approximation is 0.01. Since several fit indices offer support for our model, we conclude that the fit of the model is reasonably good and that a latent variable has been identified.

For the first period, the estimated $\beta$’s are the following.$^{14}$ In the Nordic countries, the estimate for the growth of the currency is 1.35 by a highly significant $t$-statistic of 4.14 and the estimate for the growth of the labor force participation rate is 0.82, also accompanied by a highly significant $t$-statistic 3.10. Then for the Mediterranean countries, the estimate for the growth of the currency is 4.75 with a highly significant $t$-statistic, 3.64, and the estimate for the growth of the labor force participation rate is 0.05, but insignificant with the $t$-statistic of 0.08. For the control group, the estimate for the growth of the currency is 1.45, again with a highly significant $t$-statistic, 5.56, and the estimate for the growth of the labor force participation rate is 0.33, also accompanied by a highly significant $t$-statistic, 2.86. The $\beta$’s do indeed seem to differ across countries.

Table 2 presents parameter estimates for $\alpha$’s and $\gamma$’s in the first regime. As the tests for the existence and timing of potential break points suggested (see Appendix B) that there might be a break in 1983 in the data for the Mediterranean economies, the results are reported with some reservation for the first regime. The government consumption, taxation, and the interest rate have significant direct effects on the labor force participation rate. As to the causes of the shadow economy, taxation has a significant but ambiguous effect on the shadow economy. While it has a positive effect in the Nordic countries and the rest of the world, it seems to have a negative effect in the catholic countries. Benefits have a significant and negative impact on the shadow economy in the Nordic countries and the rest of the world, but not in the Mediterranean countries. While the government transfers seem to bribe people away from the shadow economy in the Nordic countries and our control group, it has no effect on the shadow activities in the catholic countries. Government consumption has only weakly significant impact on the shadow economy. Lastly, the interest rate has a negative and significant impact on the shadow economy both in the Nordic and the Mediterranean countries as the theory suggests. The impact

$^{14}$All structural models are also accompanied with positive and statistically significant intercepts.
is, however, insignificant in our control group.

<table>
<thead>
<tr>
<th>Causes</th>
<th>PN</th>
<th>CS</th>
<th>RoW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α</td>
<td>t-stat</td>
<td>α</td>
</tr>
<tr>
<td>tax</td>
<td>0.200</td>
<td>5.10*</td>
<td>-0.101</td>
</tr>
<tr>
<td>transfers</td>
<td>-0.286</td>
<td>-4.05*</td>
<td>-0.021</td>
</tr>
<tr>
<td>gov’t cons.</td>
<td>0.16</td>
<td>1.77***</td>
<td>0.071</td>
</tr>
<tr>
<td>interest rate</td>
<td>-0.250</td>
<td>-2.04**</td>
<td>-0.218</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Direct effects on labor force</th>
<th>γ</th>
<th>t-stat</th>
<th>γ</th>
<th>t-stat</th>
<th>γ</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>by interest rate</td>
<td>0.100</td>
<td>0.57</td>
<td>0.042</td>
<td>0.19</td>
<td>-0.139</td>
<td>-2.60*</td>
</tr>
<tr>
<td>by tax</td>
<td>-0.089</td>
<td>-1.16</td>
<td>-0.117</td>
<td>-1.16</td>
<td>-0.019</td>
<td>-0.38</td>
</tr>
<tr>
<td>by gov’t cons.</td>
<td>0.127</td>
<td>1.60</td>
<td>0.160</td>
<td>1.70***</td>
<td>0.149</td>
<td>2.56**</td>
</tr>
</tbody>
</table>

Table 2: Estimates for the first regime, 1979-92 with *, ** and *** 1, 5 and 10 per cent levels of significance.

Table 3 presents the results in the latter regime. This time, the tests support the constrained model, in which \( \beta_{mi} = \beta_{mj} \), i.e., the parameters are the same across the groups.\(^{15}\) The common \( \beta_2 \), the effect of the shadow economy on the labor force is then 1.23 and is accompanied by a highly significant \( t \)-statistic, 5.88, but \( \beta_3 \), the effect on money, is 0.71 and insignificant. This time, there is only one significant direct effect - that of taxation on the labor force participation rate.

Compared with the results of the first regime, the estimator of the tax variable is positive and highly significant. We take this as an indication of the tax morale in catholic south is similar to that of the protestant north and the control group. The transfer payments have a negative and significant effect in all these groups as well. It seems a robust finding that the transfer payments tend to decrease the size of the shadow economy. Clearly, since people have to make a living in one way or another, a government concerned with the welfare of its citizens should offer benefits and transfers for those who cannot go without. Government consumption on the

\(^{15}\)Even in the unconstrained model, the following estimates for \( \alpha \)'s would be almost the same, i.e. the reported parameter changes are not caused by a change in the model definition.
other hand tends to increase shadow activity in the Nordic countries and our control group, but not in the Mediterranean countries. Does this finding suggest that people in the Nordic countries and in the control group regard government consumption as waste by the government?\textsuperscript{16} The opportunity cost of holding money has lost much of its importance in the latter period, although now also being significant for our control group. This could be an indication of that as there are now more variety for the means of payments at the official economy, the currency has gone more underground.

<table>
<thead>
<tr>
<th>Causes</th>
<th>PN</th>
<th>t-stat</th>
<th>CS</th>
<th>t-stat</th>
<th>RoW</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>tax</td>
<td>0.190</td>
<td>4.98*</td>
<td>0.200</td>
<td>4.78*</td>
<td>0.266</td>
<td>8.65*</td>
</tr>
<tr>
<td>transfers</td>
<td>-0.174</td>
<td>-3.45*</td>
<td>-0.122</td>
<td>-2.73*</td>
<td>-0.117</td>
<td>-3.92*</td>
</tr>
<tr>
<td>gov’t cons.</td>
<td>0.183</td>
<td>2.51*</td>
<td>0.017</td>
<td>0.43</td>
<td>0.132</td>
<td>4.17*</td>
</tr>
<tr>
<td>interest rate</td>
<td>-0.197</td>
<td>-2.69*</td>
<td>-0.167</td>
<td>-3.92*</td>
<td>-0.080</td>
<td>-2.03**</td>
</tr>
</tbody>
</table>

Table 3: Estimates for the second regime, 1992-2003 with *, ** and *** 1, 5 and 10 per cent levels of significance.

In a sense, our results appear to contradict those of Alm and Torgler (2005), who suggest that southern Europeans have lower tax morale than northern Europeans. While the magnitudes of the estimated coefficients of the tax variable differ in the first period they are virtually the same in the later period across the groups of countries. When it comes to the government transfers, it appears that our results are more in line with Giuso et al. (2003), who find that in the Nordic countries attitudes are more firmly against claiming government benefits when people are not entitled to them. There the government may rely more on the notion that the welfare benefits are not to be abused. Interestingly and according to our study,

\textsuperscript{16}We also tested this effect by using data on government investment. It had a similar, albeit smaller, impact on the shadow economy.
the attitudes seem to have converged, since the estimates have come closer to each other, which is in line with the findings of Feld et al. (2008) on the experiment after German unification. Perhaps during the later period the norms have become global.

3.4 Caveats

Our data is not multivariate normal, whereby the estimates for standard errors are downsized. Since we could not produce the bootstrap estimate for standard errors, we have to make inferences based on potentially fallible results. This is a serious drawback, which we could not overcome. On the positive side, we were able to demonstrate the causal variables were less susceptible to endogeneity than expected. Foremost, the indicator variables were also more independent on each other than Breusch (2005b) supposes. The reason for these findings might be the fact that we used growth rates rather than level variables.

Comparing our results across groups of countries in the two regimes, we find that there are significant differences in the parameter estimates. These differences should be taken into account when calibrating the size of the shadow economy. This problem is magnified by the existence of direct effects which, if omitted, will be carried by the indirect effects, i.e., the estimates for \( \alpha \)'s and \( \beta \)'s. While our method has its problems we show that Breusch's critique considering the direct effects might be valid. Our results demonstrate the previous studies have probably presented fallible results.

4 Final Remarks

The econometric results provide support for the proposed model, albeit some caveats remain. Taxation, government consumption and social security were all found significant. As a policy implication, this finding suggests that an increase in the size of the public sector with the high tax burden has side-effects: hidden economy expands. The significant negative estimator for the benefits points to the possibility that they might operate like a bribe on people, persuading them out of the shadow economy.
No evidence is found to support the view that tax morale today differs between the catholic south and protestant north. Moreover, since both the effect of tax and the effect of social security on shadow economy has converged across countries, there is indication on that attitudes towards public policy have converged.

With regard to the estimates for the size of the shadow economy, several important empirical results are found in this paper. We used the data typically used when quantifying the size of the shadow economy and demonstrated the data is not multivariate normal. Also the causal variables are susceptible to endogeneity as proposed by Breusch (2005b). Moreover, direct effects are significant and regime switches have taken place. Our results demonstrate the previous studies have probably presented fallible results. Subsequently, calibrating the index of the size of the shadow economy is more complicated than thought.
A Proof of Proposition 3

In the equilibrium with shadow economy, the number of tax paying entrepreneurs is \( \frac{m}{x_1} = \frac{m}{\sqrt{2(\tau + \delta)}} \). Then, the total tax revenue is

\[
T = \frac{\tau}{x_1} \left( 1 - \frac{\sqrt{2(\tau + \delta)} - \sqrt{2\delta/\beta}}{g(\alpha_1 - \alpha_2)} \right) \frac{\tau}{\sqrt{2(\tau + \delta)}}.
\]

(10)

Developing the first derivative,

\[
\frac{\partial T}{\partial \tau} = -\left( \tau - \frac{(\tau + 2\delta)}{\sqrt{2(\tau + \delta)}} \left( g(\alpha_1 - \alpha_2) - \sqrt{2(\tau + \delta)} + \sqrt{2\delta/\beta} \right) \right) \frac{1}{2(\tau + \delta) g(\alpha_1 - \alpha_2)}.
\]

The revenue maximizing tax rate \( \tau^* \) satisfies \( \frac{\partial T}{\partial \tau} = 0 \). Yet, no closed-form solution is available. However, it holds that \( T(0) = 0 \). Evaluating the first-order condition at the origin, \( \frac{\partial T}{\partial \tau}_{\tau=0} = \frac{1}{g(\alpha_1 - \alpha_2)\sqrt{2\delta}} \left( g(\alpha_1 - \alpha_2) - \sqrt{2\delta} + \sqrt{2\delta/\beta} \right) > 0 \) Thereby, the Laffer-curve is increasing at the origin. Developing the second derivative, \( \frac{\partial^2 T}{\partial \tau^2} \),

\[
-\frac{2(\tau + 2\delta)}{4g(\tau + \delta)^2(\alpha_1 - \alpha_2)} - \frac{\delta \sqrt{2} \left( g(\alpha_1 - \alpha_2) + \sqrt{2\delta/\beta} - \sqrt{2(\tau + \delta)} \right)}{4g(\tau + \delta)^2 \sqrt{\tau + \delta(\alpha_1 - \alpha_2)}} < 0,
\]

reveals that the curve must have a global maximum. In addition, we know that at \( \tau = \tau^{\text{max}} \), where

\[
\tau^{\text{max}} = \frac{1}{2} \left( g(\alpha_1 - \alpha_2) + \sqrt{2\delta/\beta} \right)^2 - \delta,
\]

the legal sector ceases to exist. Thereby, a unique revenue maximizing tax rate \( 0 < \tau^* < \tau^{\text{max}} \) must exist with \( T(\tau^*) > 0 \).

B Data

B.1 Definitions

In Table 4, we list the variables used in this study. Column 1 lists the abbreviations for the variables and Column 2 the names of the variables as they appear in the (primary) data source. The main source of data is OECD Economic Outlook 76, except for the currency in circulation. For the members of the EU and the ERM
<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>Short-term market interest rate</td>
</tr>
<tr>
<td>$x$</td>
<td>Real GDP per capita</td>
</tr>
<tr>
<td>$m$</td>
<td>Real Currency in circulation, per capita</td>
</tr>
<tr>
<td>$l$</td>
<td>Labour force participation rate</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Social Benefits Paid by Government$^{17}$</td>
</tr>
<tr>
<td>$g$</td>
<td>Government consumption</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Total direct and indirect taxes</td>
</tr>
<tr>
<td>$GDP$</td>
<td>Nominal gross domestic product</td>
</tr>
<tr>
<td>$pop$</td>
<td>Population of a country</td>
</tr>
<tr>
<td>$P$</td>
<td>Consumer price index</td>
</tr>
<tr>
<td>$s$</td>
<td>GDP deflator</td>
</tr>
</tbody>
</table>

Table 4: List of the variables and their sources.

I or II (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal and Spain) we had to retrieve data from the IMF IFS and convert the currency by using the ERM I exchange rates.

Table 5 presents the averages of each variable for both periods across countries. Reading from this table, it is generally the case that the growth rate of taxes has been greater in the first than in the second period. This is also true for the benefits variable, but not for the government consumption. Since the two activities of the government have grown faster in the first period, we name the periods as the regime of the growing public sector and of the mature public sector.

### B.2 Break Point Tests

There is casual evidence for the idea that there might be breaks in the public sector variables during the economic turbulence years of early 1990’s. Therefore, we estimate the following model by OLS for all three country groups ($t =$ calender time)

$$q_{i,t} = a_{i,t} + b_{i,t}t$$

(11)
<table>
<thead>
<tr>
<th>Country</th>
<th>$x$</th>
<th>$l$</th>
<th>$m$</th>
<th>$\tau$</th>
<th>$g$</th>
<th>$\delta$</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
<td>I</td>
</tr>
<tr>
<td>Australia</td>
<td>.57</td>
<td>1.1</td>
<td>.37</td>
<td>.21</td>
<td>1.4</td>
<td>2.1</td>
<td>4.4</td>
</tr>
<tr>
<td>Austria</td>
<td>.84</td>
<td>.67</td>
<td>-.22</td>
<td>-.06</td>
<td>.47</td>
<td>-.22</td>
<td>2.7</td>
</tr>
<tr>
<td>Belgium</td>
<td>.87</td>
<td>.68</td>
<td>.04</td>
<td>.58</td>
<td>-.13</td>
<td>.10</td>
<td>2.3</td>
</tr>
<tr>
<td>Canada</td>
<td>.46</td>
<td>.93</td>
<td>.57</td>
<td>.24</td>
<td>1.1</td>
<td>.33</td>
<td>3.82</td>
</tr>
<tr>
<td>Denmark</td>
<td>.57</td>
<td>.74</td>
<td>.39</td>
<td>-.20</td>
<td>1.7</td>
<td>1.1</td>
<td>3.5</td>
</tr>
<tr>
<td>Finland</td>
<td>.78</td>
<td>.96</td>
<td>-.01</td>
<td>-.01</td>
<td>1.2</td>
<td>3.1</td>
<td>4.0</td>
</tr>
<tr>
<td>France</td>
<td>.72</td>
<td>.58</td>
<td>-.20</td>
<td>.41</td>
<td>-.86</td>
<td>1.7</td>
<td>3.8</td>
</tr>
<tr>
<td>Germany</td>
<td>.54</td>
<td>.43</td>
<td>.56</td>
<td>.30</td>
<td>.97</td>
<td>-.13</td>
<td>2.5</td>
</tr>
<tr>
<td>Greece</td>
<td>.17</td>
<td>.89</td>
<td>.35</td>
<td>.72</td>
<td>-.11</td>
<td>.21</td>
<td>8.5</td>
</tr>
<tr>
<td>Ireland</td>
<td>1.24</td>
<td>2.55</td>
<td>-.04</td>
<td>.97</td>
<td>.12</td>
<td>1.3</td>
<td>5.0</td>
</tr>
<tr>
<td>Italy</td>
<td>.95</td>
<td>.57</td>
<td>-.10</td>
<td>.25</td>
<td>.17</td>
<td>.38</td>
<td>6.9</td>
</tr>
<tr>
<td>Japan</td>
<td>1.4</td>
<td>.40</td>
<td>.43</td>
<td>.26</td>
<td>2.0</td>
<td>1.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Netherl.</td>
<td>.65</td>
<td>.73</td>
<td>.19</td>
<td>1.0</td>
<td>.69</td>
<td>-.67</td>
<td>2.0</td>
</tr>
<tr>
<td>New Zeal.</td>
<td>.51</td>
<td>1.0</td>
<td>-.29</td>
<td>.31</td>
<td>1.2</td>
<td>2.2</td>
<td>1.5</td>
</tr>
<tr>
<td>Norway</td>
<td>1.1</td>
<td>1.1</td>
<td>.17</td>
<td>.28</td>
<td>1.3</td>
<td>1.4</td>
<td>3.6</td>
</tr>
<tr>
<td>Portugal</td>
<td>1.4</td>
<td>.72</td>
<td>.48</td>
<td>.39</td>
<td>-.12</td>
<td>1.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Spain</td>
<td>.88</td>
<td>.93</td>
<td>-.22</td>
<td>1.35</td>
<td>1.4</td>
<td>-.57</td>
<td>7.2</td>
</tr>
<tr>
<td>Sweden</td>
<td>.63</td>
<td>.78</td>
<td>.05</td>
<td>-.56</td>
<td>-.14</td>
<td>1.1</td>
<td>3.9</td>
</tr>
<tr>
<td>Switzerl.</td>
<td>.59</td>
<td>.24</td>
<td>1.00</td>
<td>-.09</td>
<td>.44</td>
<td>.69</td>
<td>1.6</td>
</tr>
<tr>
<td>UK</td>
<td>.65</td>
<td>1.0</td>
<td>.25</td>
<td>.08</td>
<td>2.4</td>
<td>1.5</td>
<td>4.1</td>
</tr>
<tr>
<td>the US</td>
<td>.75</td>
<td>0.83</td>
<td>.37</td>
<td>.01</td>
<td>.05</td>
<td>.61</td>
<td>2.9</td>
</tr>
<tr>
<td>Average</td>
<td>.77</td>
<td>.85</td>
<td>.19</td>
<td>.31</td>
<td>.57</td>
<td>.91</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Table 5: Averages of each variables over the first and second period, as per cents.
and test for the presence of a break, examining whether $a_{i,t} = a_{i,t-1}$ and $b_{i,t} = b_{i,t-1}$ for all $t$. The Quandt-Andrews unknown breakpoint test seeks the presence of a breakpoint at any $t$ and reports the maximum of LR-statistic and the year of the breakpoint. The results are presented in Table 6.

The results indicate that both the benefits in the North and in the Rest of the World exhibited a break in 1992 while the tax variables exhibited the break already in 1990. There is no evidence of a break in the government expenditures in these two groups. The results for the Catholic South suggest a break in 1983 both in the benefits and in the government expenditure variables while the break in the tax variable appears a year earlier. Our findings suggest that multiple breaks are present. Yet to have at least one period where no breaks occur, we choose to split our data in 1992.

### C Endogeneity

Typically the endogeneity of the instruments can be tested by Granger-causality test, whereby "$y$ should not cause $x$". While this test is easy to implement, the problem is that it does not control for other variables besides one proposed cause. This may result wrong inferences on the endogeneity of the variable $x$, due to missing variable problem. Using the more general version, whereby "$x$ should be independent of past $y$ conditional on past $x$" is probably closer to the true model and mitigates the problem of missing variables. Therefore to test the endogeneity
of the variables we test the following two propositions

\[ x_{j,t} = \beta_0 + \sum_{k=1}^{2} \beta_{j,t-k}x_{j,t-k} + \sum_{i=1}^{3} \sum_{k=1}^{2} \beta_{i,t-k}y_{i,t-k} + \epsilon_{j,t}, \quad (12) \]

where \( x_{j,t} \) is the current value of each cause variable \( j \) and \( y_{i,t-k} \) are the past values of each indicator variables \( i \). As our \( T = 25 \) in the full sample and since we work with annual data, we set \( k = 2 \), i.e. we have two lags of each variable in the model we test.\(^\text{18}\)

It is also crucial that indicator variables are independent on each other. Testing it in the presence of non-observed variable, which we expect to affect all indicators is, however, difficult. What one can do is to propose a model where the non-observed variable is replaced by the cause variables as in the instrument variables procedure to control for latent variable’s absence from the model. Therefore we proceed to test the following model

\[ y_{j,t} = \beta_0 + \sum_{i=1}^{3} \sum_{k=1}^{2} \beta_{i,t-k}y_{i,t-k} + \sum_{j=1}^{4} \beta_{j,t-2}x_{j,t-2} + \epsilon_{j,t}, \quad (13) \]

where the current values of each indicator variable \( y_{j,t} \) are explained by the lags of all three indicators and the the lags of all proposed causes. In order to save the degrees of freedoms we set the number of lags for indicators to 2 and include only the second lag of each proposed cause variable.

We first test the simple Granger causality applied to the pooled, cross-country data of our cause variables. The Granger causality test suggests that our measure of tax is probably caused by real GDP per capita, which seems plausible. The transfers, on the other hand, might be caused by labor force participation and real GDP per capita - again a result in accordance with the expectations. Government consumption might be caused by real GDP per capita and interest rates are caused both by real GDP per capita and currency in circulation. As to indicators, labor force participation might be caused by real GDP per capita and somewhat surprisingly might real GDP per capita be caused by real currency in circulation, but not the other way around. According to Granger causality tests, there is some evidence that the proposed causes and indicators are not exogenous.\(^\text{19}\)

\(^{18}\)Other lag lengths were also tested and the results were similar to those reported here.

\(^{19}\)All the results presented here are available from the author at request.

84
Table 7: Results from the general causality test for proposed cause variables *, ** and ***, 1, 5 and 10 per cent levels of significance.

Table 8 presents the results for indicator variables. Since we are not interested on the effects of our causes for the indicators per se, we suppress the information and present whether there was any significant cause for the indicators. As it turns out, the second lags of our proposed cause variables seem not to cause real currency per capita. The more interesting result is that, controlling for the cause variables, our indicators seem to be independent on each other. Only the labor force participation seems to be affected by the first lag of real GDP per capita. Since the bias introduced by the fact that the indicator variables are not exogenous is far greater than the
bias due to endogeneity of the cause variables, this is a good sign.

To summarize, the endogeneity of our cause and indicator variables is less of a problem than one would have expected. While our results are probably not heavily influenced by the endogeneity bias, the estimated standard errors are instead downsized because of non-normality and small sample size.

## D Fit criteria

In principle there are three criteria according to which one may judge the significance or goodness of the model. The first and the obvious one is the magnitude and the direction of the parameter estimates and whether they are in accordance with underlying theory. The second, also related to the first one, is the statistical significance of individual parameter estimates, measured typically by the \(t\)-test. The third criteria are the global fit measures.

There are several fit criteria available for the SEM users, although they seem to be closely related to each other. \(\chi^2\)-test compares two models; the saturated, theoretical model, which includes all paths and the hypothesized model we impose. Since the idea is to find an optimal parsimonious presentation on the true structures
in the data, the simplest model that conveys the information in the data would be ideal. A low and statistically insignificant $\chi^2$-value suggests that our model resembles that of the saturated model. However, the $\chi^2$-test statistic is affected by several factors, such as the sample and the model sizes, i.e., bigger the model the better the fit. Hence we also report other fit indices, albeit to some extent they suffer from the same problems as they are derived from $\chi^2$-tests.\footnote{AMOS 17 reports long list of fit indices. Since some of them are just new representations of the older fit indices, we report only those that seem to offer some new insights to the analysis.}

The normed fit index (NFI) is computed by comparing the $\chi^2$-value of the null model (i.e. independence model, where the covariances are assumed to be zero in the model) and the hypothesized model, whereby $NFI = (\chi^2_{null} - \chi^2_{model}) / \chi^2_{null}$. The idea is to test how much better our hypothesized model performs compared with the null model. Conventionally a test value close to 0.95 indicates a good fit. The relative fit index is similar to NFI, but it accounts for the different decrees of freedom these two models have, whereby $RFI = 1 - [(\chi^2_{model} / df_{model}) / (\chi^2_{null} / df_{null})]$. We also report Bentler’s comparative fit index (CFI), and Tucker-Lewis Index (TLI) which are written as

\[
CFI = 1 - \frac{(\chi^2_{model} - df_{model})}{(\chi^2_{null} - df_{null})} = 1 - \frac{\lambda_M}{\lambda_N},
\]

\[
TLI = \frac{\chi^2_{null} / df_{null} - \chi^2_{model} / df_{model}}{\chi^2_{null} / df_{null} - 1} = 1 - \frac{\lambda_M / df_M}{\lambda_N / df_N},
\]

where $\lambda_M = (\chi^2_{model} - df_{model})$ and $\lambda_N = (\chi^2_{null} - df_{null})$. While CFI and TLI seem to test the same hypothesis, i.e. they test the null of no latent variable against the alternative of latent variables existing, the former is affected by the sample size as well as non-normality of the data (Bollen, 1990). TLI, in the other hand, has proven to be least affected by the biasing factors (non-normality, small sample size, model complexity etc.).

Lastly to avoid the problems documented with the fit criteria related to $\chi^2$-test we report standardized root mean square residual (SRMR). The test reflects the closeness of the empirical covariance-variance matrix ($S$) and the model implied covariance matrix ($\Sigma$). Comparing these two matrices reveals whether all the connections present in the data are presented in the model. The better the model
reflects the data smaller the SRMR statistics. While there is no rule of thumb for the cut-off value, values smaller than 0.05 are acceptable.
References


1 Introduction

The 1991 collapse of the Soviet Union created 15 new independent states. These, along with the other countries in Eastern Europe, entered in a transition from centrally planned economy to market economy. Each met with varying degrees of success; some posted solid growth, others struggled with sharp reversals of fortune (see Havrylyshyn and Wolf, 1999). This paper analyzes the determinants of growth in transition. In line with the emerging academic consensus, we present evidence that successful governmental and institutional reforms are necessary conditions for sustained growth.

The debate on determinants of growth initially crystallized around Solow’s seminal 1956 paper. As growth studies evolved, there emerged a recognition that poor protection of property rights impairs growth by reducing incentives to invest (Mauro, 1995). The research community a half century later now generally concurs that good government is critical for economic success (e.g. Acemoglu, 2008; Giavazzi and Tabellini, 2005), since investment and technological advances can be easily disturbed by bureaucratic propensities to rent-seeking or corruption. Indeed, the lousy economic performances of transition economies were soon linked to their institutional shortcomings (e.g. Frye and Shleifer, 1997).

Several papers seek to determine the effect of institutions on growth in transition. Fidrmuc and Tichit (2009) suggest that the data is vulnerable to structural breaks across time and/or countries. They note that the pattern of growth in transition has changed at least twice; yielding three distinct models of growth associated with different stages of reform. The third regime started in mid 1990s. Babetskii and Campos (2007) conduct meta-analysis to investigate the effect of institutional reform to growth. They find that approximately a third of papers find a positive and significant relationship, another third finds a negative and significant relationship, and a third find no significant relationship between reform and growth.

Institutions and growth might be jointly determined, whereby an exogenous proxy for institutions is needed. Glaeser et al. (2004) completely reject the argument that institutions cause growth, claiming the causation actually works in the opposite direction; i.e. growth and human capital accumulation drive institutional development. They further provide evidence that human capital rather than po-
litical institutions is the crucial element of growth. Galor et al. (2008) extend this argument with their model treating human capital promoting institutions as primary to growth.

In the present discussion, we empirically investigate the relationship of the government, institutions, human capital, and economic growth. Transition economies form an ideal set for study as they have all been part of a natural experiment. Due to communism, they share relatively similar history. Moreover, all faced the same shock as they abandoned communism and command economies; all inherited dysfunctional institutions. This shock and the following structural change causes a break between the level of development (growth) and institutions as proposed by Glaeser et al. (2004). The citizens in all these countries are generally well educated, and perhaps more important for our purposes here, education levels, literacy rates, etc. were similar across this group at the start of transition. The relative similarity of human capital stock allows us to examine for differences as they emerge across countries as transition progresses, particularly with respect to reforms of economic and political institutions.\footnote{Åslund (2007) proposes that when the former communist block was reformed in early 1990s, liberal reformers won out in Central Europe and the Baltics, while rent-seekers came to dominate in CIS countries.} It also provides an opportunity to distinguish the effect of institutions on growth from the effect of human capital since as we show later the latter has been almost stable across time and space evaluated in this paper.

Although many studies acknowledge that informal (or illegal) production accounts for a significant chunk of total production in transition countries, most base their analysis solely on official output growth figures. This is a huge omission. For example, Schneider (2004) estimates the unweighted average of the size of the shadow economy in transition economies during 2002–2003 equaled 40.1% of official GDP, implying that nearly 30% of total production in transition economies occurred underground. Feige and Urban (2008), on the other hand, note the weaknesses of measurements of underground activities. They propose that conclusions concerning the success of transition rely heavily on recorded measures of GDP and thus viewed with skepticism. We attempt to correct for the omission of underground production by evaluating the growth of real GDP per worker, a measurement of
productivity. By concentrating on those listed as employed in the formal economy, GDP per worker gives a sharper picture of average productivity and the growth potential of these nations.

We use yearly data from 1998 to 2005 to avoid possible breaks present in the early years of transition in these 25 economies, and control for endogeneity using the dynamic GMM method proposed by Arellano and Bond (1991). As the problems of the two-step GMM are generally well-documented (see Windmeijer, 2005 and Roodman, 2009), we use the proposed measures to correct the biases in this method. Applying a regression tree analysis, we test for regimes of growth (i.e. non-linearity) with respect to human capital or institutions. While testing for regimes of growth, we use a fairly rich set of variables, for example, we recalculate the Human Development Index, excluding GDP per capita, to portray the evolution of human capital. Also following Fidrmuc and Tichit (2009), we calculate a weighted average of transition indicators to proxy the evolution of institutions. We test several interaction terms to allow for non-linearity in the growth model we estimate.

Our findings contain several notable insights. First, we find no regimes on growth; all countries surveyed obey the same model and laws of motion and the relationship between growth and human capital (institutions) is linear across space. Second, at the beginning economic freedom and investment contribute positively on growth in transition. When the countries reach their population averages in terms of economic freedom or investment the negative interaction takes its toll; the full marginal effect of investment (economic freedom) to growth is negative if the economic freedom is above its average. Third, increased government consumption (our measure of the size of public sector) seems to have a negative impact on growth. These findings are robust when we drop resource-rich countries from the dataset. Finally, our robustness analysis shows similar results when evaluating real growth per capita, although investment and size of government seem to matter more for productivity growth than growth per capita.

Comparing previous estimations of growth in transition economies against our results vindicates a number of research claims. We confirm the results of Fidrmuc and Tichit (2009), who suggest that the countries surveyed adhered to a common growth model during the later years of transition. Contrary to some earlier findings,
our results infer that different government policies affect growth in terms of forming institutions and setting the size of the government. How we measure institutions apparently makes a difference; our results change when we use Fidrmuc and Tichit’s recalibrated index of economic reform. Moreover, when measuring productivity and economic well-being, it makes a difference whether one uses real GDP per capita or per worker. Our results indicate some of the contradictory results in the earlier literature may arise from the use of an inappropriate model, yet we also confirm many previous findings.

The paper is structured as follows. Section 2 reviews the growth theory literature, particularly key empirical studies. Those familiar with the literature can skip the review and go straight to Section 3 for a presentation of the empirical results. Section 4 concludes.

2 Literature

Soon after Barro (1990) published his frame-breaking work modelling public services in an endogenous growth setting, he was followed with an extension of the analysis to capture varieties of public goods (Barro and Sala-i-Martin, 1992). Devarajan et al. (1996) responded with a model postulating two types of government expenditure, productive and unproductive, to show how changes in the composition of public expenditure affect long-term economic growth rates.

Following a different track, Lucas (1988) and Mankiw et al. (1992) proposed that the human capital accumulation is essential for economic growth. Hall and Jones (1999) synthesized these perspective into a model in which social infrastructure influences growth via production inputs. Recently Galor et al. (2008) extend these arguments with their model treating human capital promoting institutions as primary to growth.

Throughout the literature, there seems to be a common acceptance of the notion that bad economic policies harm development. The political economy view emphasizes the role of the market reforms and controlling against opportunities for rent-seeking and corruption. Vested interests for government officials can halt the reform process, create incentives for the underground production, and slow progress
in the official economy (Harstad and Svensson, 2006). Moreover, the role of institutions has been emphasized and disputed, as several papers seek to determine the effect of institutions on growth. According to Acemoglu et al. (2001, 2002), incomes improved in colonies where Europeans developed good institutions. The institutions hypothesis has also gained empirical support: Easterly and Levine (2003) and Esfahani and Ramirez (2003) in the area of contract enforcement, Djankov et al. (2006) in the area of business regulations, and Acemoglu and Johnson (2003a) and Brunt (2007) in the area of property rights. As noted by Hanushek and Woessmann (2008), education and skills may not have the desired impact on economic outcomes in the absence of proper institutions.

It is often suggested that public goods are a vital part of the private production due to *strategic complementarity*, whereby the composition of public expenditure affects growth. Devarajan et al. (1996) find that the share of current expenditure has a positive effect on growth, while an increase in the capital component of public expenditure has a negative growth effect. They characterize this as the “too much of a good thing” effect of capital expenditure, i.e. excessive spending in developing countries renders them unproductive at the margin. Shioji (2001) finds that the infrastructure component of public capital has a significant and positive effect on long-run output in the US and in Japan. Blankenau et al. (2007) find a positive relationship between public education expenditures and long-term growth after controlling for government budget constraints. Aschauer (2000) argues that the relationship between public capital and economic growth is non-linear. Also Minier (2007) tests the non-linearities of growth model and shows that either squared terms or interaction terms of fiscal variables should be present in the model. Moreover, allowing for non-linearities, several fiscal variables become robust.

---

2 For an excellent debate and theoretical analysis, see Acemoglu (2008), who contrasts the oligarchic and democratic societies and studies the entry barriers in place. Under oligarchic regimes, the elite withhold the monopoly position. In democracies, taxes create distortions. See also Frye and Shleifer (1997) and their grabbing-hand model. For further discussion, see La Porta et al. (1999).

3 This view is further elaborated by Acemoglu et al. (2003b), who show that distortionary macroeconomic policies are more likely to be symptoms of underlying institutional problems. Fogli (2003) presents a critical view, proposing that technological adoption is significantly linked to institutional variables and that its omission is not neutral to the analysis.
There have been other attempts to define the determinants of growth and reduce model uncertainty. Durlauf et al. (2005) list 145 potential explanatory variables in growth regressions. Magnus et al. (2008) attempt to reduce model uncertainty and determine the “focus” and “auxiliary” regressors for growth. They find that constant, initial GDP per capita, real equipment investment share of GDP, initial total gross enrollment ratio for primary education, and life expectancy at age 0 are the focus variables. Thereafter, average growth rate of population, rule of law, tropical land area, ethnolinguistic fragmentation, and fraction of Confucian population (as a proxy for religion or culture) are found to be focus or auxiliary variables depending on the model.

3 Growth in Transition

3.1 Measurement

When measuring the income of nations, standard real GDP per capita can be misleading. As the output of the informal sector is not directly measurable, the official GDP per capita figure likely underestimates the true prosperity of a country with a sizeable informal sector. For example, Schneider (2004) finds that the estimated size of the shadow economy in transition economies during 2002–2003 ranged from 20.1 in the Czech Republic to 68.0 in Georgia. In terms of total production (legal plus illegal), this means that “non-observed” production ranged from 16.7% to 40.5% of the total. Thus, welfare comparisons and productivity growth estimates may be misleading if we concentrate solely on measured production.

One possible way to correct for this measurement error is to use GDP per worker instead of GDP per capita. Those not part of official employment are likely to make their living outside the formal economy, so real GDP per worker tracks total productivity much more closely.\(^4\) Hence, the evolution of real GDP per worker should offer insight into actual wealth and growth potential of these nations. We

\(^4\)Using GDP per worker has an additional benefit. It is a closer measure of standards of living as it corrects for home production and leisure.
illustrate this as

\[
\frac{Y}{POP} = \frac{Y}{L} \cdot \frac{L}{POP}
\]

where \(Y\) is the real income, \(L\) is the labor force and \(POP\) is the population. Equation (1) suggests that real GDP per capita can rise either from growth in labor productivity or because labor force participation increases. Here, we concentrate on the former.

To estimate the relationship between real GDP per worker and institutions, the following model is proposed

\[
y_{t;i} = \gamma + \gamma INS_{t;i} + \beta X_{t;i} + \mu_{t;i},
\]

where \(y_{t;i}\) is the log of real GDP per worker, \(INS_{t;i}\) is the measure of the quality of the institutions and \(X_{t;i}\) are the control variables. Since \(y_{t-1;i}\) is correlated with \(\mu_{t;i} + v_{t;i}\), the use of an instrumental variables approach is preferred. Here, we use Arellano and Bond’s dynamic GMM. Subtracting the lagged version of (2) from \(y_{t;i}\) we obtain

\[
y_{t;i} - y_{t-1;i} = \delta (y_{t-1;i} - y_{t-2;i}) + \beta (X_{t;i} - X_{t-1;i}) + \gamma (INS_{t;i} - INS_{t-1;i}) + \mu_{t;i} + \nu_{t;i},
\]

where twice or more lagged levels of the dependent variable and predetermined variables can be used as instruments for (3).

It has been found that the estimated asymptotic standard errors of the Arellano and Bond’s two-step estimator exhibit downward bias in small samples (see Windmeijer (2005)). Moreover, since the internal instrument set includes the past observations of the instrumented variables, the number of instruments grows quadratically with respect to \(T\). According to Roodman (2009), a large collection of (weak) instruments overfits endogenous variables and weakens the Hansen test of the instruments’ joint validity. While the practice has often been to reduce the number of instruments by using only certain lags, Roodman (2009) recommends “collapsing” either alone or with constrained lags. This way the instrument count will be linear in \(T\), which mitigates the problem of overfitting bias. With respect to bias in standard errors, Windmeijer (2005) provides a method to correct the estimate of the variance. He also provides evidence for that the corrected two-step GMM
estimator outperforms the uncorrected two-step estimator and may even outperform the one-step estimator. We use these as they are options available in STATA (xtabond2).

The annual data for 25 transition economies is drawn from multiple sources, since none of them provides a complete set of data for all countries even over our short time horizon (1995-2005). Real GDP per worker data are drawn from World Bank World Development Indicators, while the share of government expenditure relative to GDP is taken from the European Bank of Reconstruction and Development SEI database. Last, the Penn World Tables (version 6.2) provide data on the share of investment relative to GDP (for details, see Appendix A).

To measure human capital, we calculate a re-scaled human development index (HDI), basically an average of life expectancy index and education index as reported in the UN’s Human Development Report. We interpret it as a broad measure of the human capital accumulation (see Kalaitzidakis et al., 2001). For the measure of economic freedom, we use the Heritage Foundation’s Index of Economic Freedom (EFI). To compare our results against other studies, we also calculate the weighted average of transition indicators (EBRD). Thus is determined using weights proposed by Fidrmuc and Tichit (2009), who conduct a factor analysis to determine how well each EBRD indicator reflect progress in transition. Our final measure of transition is thus the composite of eight indicators reported by the EBRD rather than the overall unweighted indicator reported by the EBRD.

5Originally HDI has also an third indicator as a measure of decent standard of living, namely GDP per capita re-scaled to a GDP index. Two remaining indices are life-expectancy at birth (life-expectancy index) and education index. The latter is comprised of adult literacy index (adult literacy rate) and GER index (gross school enrolment ratio).

6The Heritage Foundation’s Index of Economic Freedom comprises 10 components: business freedom, trade freedom, fiscal freedom, government size, monetary freedom, investment freedom, financial freedom, property rights, freedom from corruption, and labor freedom. For detailed information on the each sub-category, go to www.heritage.org.

7These 8 indicators reflect the scale of reform in different sectors of an economy. The indicators and the assigned weights are as follows: price liberalization (0.08), foreign exchange and trade liberalization (0.15), small privatization (0.13), large privatization (0.14), enterprise reform (0.21), competition liberalization (0.10), bank liberalization (0.18) and non bank liberalization (0.09). For more details see Fidrmuc and Tichit (2009) and the EBRD Transition Indicators Database.
To verify the hypotheses that these countries obey the same growth model and that there are no non-linearities with regard to human capital or institutions, we first determine if the data can be divided into “clubs” using the regression tree method proposed by Durlauf and Johnson (1995). Calculate the average growth of real GDP per worker (1995-2005), and applying indicators for the initial (i.e. 1995) level of human capital and institutions, their changes and average growth rates, we find there is no efficient split with regard to any of aforementioned variables. This finding suggests that during the observation period these countries obey either the same laws of motion, or at least relatively similar growth models as proposed by Fidrmuc and Tichit (2009). This is reasonable since the countries share fairly similar histories.

Figure 1 in Appendix B presents the 1995 and 2005 values of human capital across these economies. It confirms our prior assumption that, in terms of human capital, post-communist countries were relatively homogenous at the start of transition. With respect to changes in human capital, the levels of 2005 vary more than the levels of 1995. It is safe, however, to conclude that the level of human capital has been stable across time and space. Figure 2 in Appendix B tells a different story. The variations in economic freedom across economies increase between 1995 and 2005. In this respect, human capital does not drive differences in institutions or economic growth. The transition from the communism to market economies and the break in the growth model proposed by Fidrmuc and Tichit (2009) also suggest that pre-transition levels of development are not determinative of the route or institutions adopted in the new model. The structural break in the system should break the link from the level of development to institutions.

In what follows we use share of investment, share of government consumption and two controls for the institutions namely the Heritage Foundation’s EFI and re-scaled EBRD indicator on transition as drivers of growth of real GDP per worker. As mentioned, human capital is relatively constant over time and space so it cannot be used as an explanatory variable in the estimations. Therefore, the lower diagonal of Table 1 presents the pair-wise correlations between the variables used in the

---

8To make sure that our results are not completely driven by the use of these two indicators, we also test fertility rate, Gini coefficients and overall upper secondary enrollment rates.
following estimations. Generally, the correlation is either not present or very low. The exception is the correlation between measurements of institutions (0.70). In essence, they reflect the same phenomenon and are driven by the same process. The basic statistics in terms of means and standard errors are presented in Appendix A.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Growth</th>
<th>Investment</th>
<th>Gov’t Cons.</th>
<th>EFI</th>
<th>EBRD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment</td>
<td>0.06</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gov’t Cons.</td>
<td>-0.25</td>
<td>0.05</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EFI</td>
<td>0.26</td>
<td>0.13</td>
<td>0.32</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EBRD</td>
<td>0.21</td>
<td>0.22</td>
<td>0.35</td>
<td>0.70</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Correlation among key variables.

3.2 Dynamic GMM

Our second step involves estimating the dynamic growth regression by GMM using the method proposed by Arellano and Bond (1991). The panel unit root tests generally suggest that growth of real GDP per worker is stationary. To obtain consistent estimators, the lagged levels (or differences) of the proposed explanatory variables may be used as long as they are correlated with the variable they are supposed to instrument. Evaluating the correlation with $\Delta x_t$ and $x_{t-p}$ (or $x_t$ and $\Delta x_{t-p}$ for system-GMM) for $p = 1, 2, 3$ shows that all the first differences of explanatory variables correlate up to at least their third-level lag.

The following instruments were used. Lags two and three are included for growth, and, following Roodman (2009), collapsed. For other variables, which are likely predetermined rather than strongly exogenous, only third lags are included. The instrument list is also collapsed. Standard errors, which should be consistent in the presence of any pattern of heteroskedasticity and autocorrelation within panels, are corrected for the small sample bias in line with Windmeijer (2005).

Table 2 presents the estimation results for the economic growth with different model specifications. The results vary considerably depending on the model specification used. The Sargan, and in particular, Hansen test-statistics are generally
insignificant, i.e. the instruments are valid. Moreover, the models seem to pass the Arellano-Bond test for second-order serial correlation and the Wald test for joint significance for coefficients.

The results of the baseline model, where growth is explained by its own lag, investment, government consumption, and economic freedom, show that previous growth rate and economic freedom are statistically significant while investment and government consumption are not. This finding is in line with similar earlier studies. The coefficient for lagged growth is 0.287 and highly significant (1% level of significance), while the measure of the economic freedom also obtains a statistically significant coefficient 0.01. The Wald test for joint significance for coefficients is 214.68, indicating the model is valid.

Following Minier (2007), we test several interaction terms. In particular, we introduce the interactions of economic freedom with other variables into our model. The second model brings the interaction between economic freedom and investment to the estimated model, which changes the results compared to our baseline model. The coefficient of growth is reduced, but remains significant. Both investment and government consumption now seem to have an impact on growth. The impact of investment is positive, while the impact of government consumption is negative, possibly an indication of wasteful spending. The most interesting result is that the interaction term is negative and significant: i.e. in the presence of greater economic freedom, previous investment tends to have a negative effect. We next test the interaction between economic freedom and government consumption in the third model. There are again notable changes. The baseline results reappear, but investment is now also statistically significant. Lastly, introducing the interaction between economic freedom and growth into the model makes a notable change compared to previous models as lag of growth is no longer significant.

Since all the models seem to pass the tests for measuring the performance of instruments and/or the model, it is difficult to judge which one is preferable. If the fourth model represents the full (general) model, then the last interaction terms which are statistically insignificant do not seem to bring any value-added to the

---

9 Other interactions were tested, but only those that seemed to have an impact are presented here.
model or the results. Therefore, model two with one interaction term would seem to specify a correct model. These results also have an intuitive appeal.

In order to ease the interpretation of the results, we re-estimate model two using standardized variables. First, increases in investment have a higher impact than that of any other variable with the exception of interaction. This runs counter to some previous findings. As the interaction is significant, the entire marginal effect should be considered. For the sake of simplicity, we assume we are one standard deviation above the expected level of economic freedom. Increasing investment by one standard deviation slightly decreases growth (1.6-1.7=-0.1). Moreover, assuming we are one standard deviation above the expected level of investment, we find that increasing economic freedom by one standard deviation decreases growth quite substantially (0.7-1.7=-1.0). If we are below the expected levels, the impact is the opposite: the presence of interaction increases the total marginal effect and the impact on growth is magnified.

In summary, economic freedom and investment initially impact positively on growth, whereas the size of government seems to have an negative impact on growth. However, in the presence of high level of economic freedom (investment), the impact of investment (economic freedom) seems to have a detrimental effect on growth. If we have economic freedom (investment) below or close to its expected value, then increases in investment (economic freedom) surely increase growth. There appears to be a “too much of a good thing” phenomenon at work at their joint presence. Increases in investment and economic freedom are good as long as they are not overdone.

\[\text{We also tested this idea by including the square of size of government, investments and economic freedom instead of their interactions, which all turned out to be insignificant. While the interpretation of interactions term as an indicator of a "too much of a good thing" phenomenon might seem unconventional, it is in fact what happens here. Since the total marginal effect of either one of variables becomes negative if we increase the other beyond its mean value, taken together these positive forces become detrimental.}\]
<table>
<thead>
<tr>
<th>Variables</th>
<th>Growth of real GDP per worker 1998-2005, N=25, T=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>Coeff.</td>
</tr>
<tr>
<td>(first lags)</td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>.287***</td>
</tr>
<tr>
<td></td>
<td>(.110)</td>
</tr>
<tr>
<td>Investment</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
</tr>
<tr>
<td>Gov’t Cons.</td>
<td>-.001</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
</tr>
<tr>
<td>EFI</td>
<td>.001**</td>
</tr>
<tr>
<td></td>
<td>(-6.5e^-4)</td>
</tr>
<tr>
<td>EFI*Investment</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(5.6e^-5)</td>
</tr>
<tr>
<td>EFI*Gov’t Cons.</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>EFI*Growth</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Instruments | 9 | 11 | 11 | 13 | 15 |
| Observations | 232 | 232 | 232 | 232 | 232 |
| Sargan J-test | 7.50 | 5.00 | 5.02 | 5.56 | 7.81 |
| p-value | 0.19 | 0.54 | 0.54 | 0.59 | 0.45 |
| Hansen J-test | 4.66 | 2.50 | 2.18 | 4.78 | 6.92 |
| p-value | 0.46 | 0.87 | 0.90 | 0.69 | 0.54 |
| $m_2$ | -0.80 | -0.96 | -0.79 | -0.91 | -0.62 |
| p-value | 0.43 | 0.338 | 0.43 | 0.365 | 0.53 |
| Wald test | 214.68*** | 113.4*** | 74.44*** | 346.31*** | 387.85*** |
| p-value | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Table 2: Estimation results with Arellano and Bond two-step GMM. (***) and (**) and (*) indicate that the coefficient is significant at 1, 5 and 10 % level of significance. Robust standard errors are reported below the coefficients.
3.3 Robustness Check

We perform three tests to check the robustness of the results (Table 3). In the first check, we drop the resource-rich countries from the data,\(^{11}\) and redo the analysis concentrating on non-oil economies. Surprisingly, there is little change in the results. The most notable change for non-oil economies is that lagged growth is no longer statistically significant. With respect to the general model, the results resemble those of a full sample. Turning to growth per capita estimates, we see that they appear very similar to those for growth per worker, but there are small differences in the magnitudes of the coefficients. For example, the coefficient for the investment has decreased from 0.011 to 0.007. It appears as most of the variables have smaller impact on measured production than on total production with the exception of previous growth, which now seems to have a larger impact on today’s growth. All these findings disappear, of course, when we evaluate the general version of the model; there is only one statistically significant effect in the full model. Finally, we use the recalculated index of economic reform as in Fidrmuc and Tichit (2009) instead of economic freedom. For these countries, the economic reform has strong, positive impact on growth, although most of growth seems to be explained by its own lag. The interaction is no longer statistically significant. The inference for the full model is not available, however, since the variables seem to suffer from multicollinearity that introduces a singular covariance matrix.

In summary, it makes a difference how the institutions are measured. Growth per worker and growth per capita are not explained by the same model.

Our results are in line with Fidrmuc and Tichit (2009), who find a positive relationship between growth and institutions in their third growth regime. These results are consistent with the findings of Iradian (2007), who concludes that the growth impetus associated with market reforms in CIS has been substantial due to its effect on overall productivity. Our results, however, suggest one can overdo both investment and reform. Moreover, we find a statistically significant interaction between economic freedom and investment, which might help to explain earlier, contradictory findings. As is shown in the first model, if this effect is left uncontrolled, we find no significant relationship between growth and investment. Finally, these results

\(^{11}\)Azerbaijan, Kazakhstan and Russia.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Non-oil, N=21</th>
<th>Growth per capita</th>
<th>Growth per worker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>.221</td>
<td>.299***</td>
<td>.368</td>
</tr>
<tr>
<td></td>
<td>(.139)</td>
<td>(.099)</td>
<td>(.471)</td>
</tr>
<tr>
<td>Investment</td>
<td>.014***</td>
<td>.007**</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>(.004)</td>
<td>(.004)</td>
</tr>
<tr>
<td>Gov’t Cons.</td>
<td>-.004***</td>
<td>-.007**</td>
<td>-4.7e-4</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td>(.001)</td>
<td>(.001)</td>
</tr>
<tr>
<td>EFI</td>
<td>.003***</td>
<td>.002**</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td>(5.3e-4)</td>
<td>(.001)</td>
</tr>
<tr>
<td>EFI*Investment</td>
<td>-2.2e-4****</td>
<td>-1.2e-4**</td>
<td>-6.0e-5</td>
</tr>
<tr>
<td></td>
<td>(5.3e-5)</td>
<td>(4.7e-5)</td>
<td>(8.4e-5)</td>
</tr>
<tr>
<td>EFI*Gov’t Cons.</td>
<td>-</td>
<td>-8.5e-5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(5.5e-5)</td>
<td>(7.3e-5)</td>
<td></td>
</tr>
<tr>
<td>EFI*Growth</td>
<td>-</td>
<td>-4.9e-4</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>(5.5e-5)</td>
<td>(.008)</td>
<td>(.012)</td>
</tr>
</tbody>
</table>

| Instruments       | 11            | 15                | 11                | 15               | 11            |
| Observations      | 205           | 205               | 228               | 228              | 232           |
| Sargan J-test     | 2.87          | 8.74              | 8.78              | 9.56             | 10.93         |
|                   | 0.84          | 0.36              | 0.19              | 0.30             | 0.09          |
| Hansen J-test     | 2.54          | 7.21              | 5.98              | 5.16             | 6.31          |
|                   | 0.86          | 0.51              | 0.57              | 0.74             | 0.39          |
| m_2               | -1.35         | -0.77             | 0.22              | -0.62            | -0.98         |
|                   | 0.18          | 0.44              | 0.83              | 0.54             | 0.33          |
| Wald test         | 193.18***     | 474.27***         | 130.45***         | 292.51***        | 151.82***     |
|                   | 0.00          | 0.00              | 0.00              | 0.00             | 0.00          |

Table 3: Estimation results with Arellano and Bond two-step GMM. (***), (**), and (*) indicate that the coefficient is significant at 1, 5 and 10 % level of significance. Robust standard errors are reported below the coefficients.
speak to the question on errors-in-variables in the measure of total productivity and reform posed by Babetskii and Campos (2007). How economic performance is measured makes a difference, even if real GDP per capita grows very much in line with total productivity. When using the composite index of economic freedom as a proxy for institutions, it seems to make relatively little difference if oil-rich countries are included in the dataset.

4 Conclusions

We considered here whether growth is driven by institutions and investment, concentrating on the recent experiences of the transition economies of the former Soviet Union and Eastern Europe. These countries all experienced the upheaval of moving from a planned economy to a market economy, and all started the transition process with fairly similar human capital endowments. We conclude that with a relatively skilled labor pool, and as long as there are insufficient institutions or private capital, improvements in institutions and investment tend to boost productivity growth.

Policy-makers can glean several conclusions from this study. First, in the presence of a high level of human capital as characterizes these transition economies, economic freedom, institutions and private investment do promote growth. As economic freedom (investment) increases, however, greater investment (economic freedom) appears to have a detrimental impact on growth. Our results seem to suggest that one can overdo both Investment and marketization, although with such a limited experience of transition process these conclusions might be premature.

Second, government consumption has a negative impact on growth. Since this can also be viewed as a measure of size of the government or the extent of public interventions in markets, our results seems to indicate that hyper-active government and the wasteful spending it engenders is a hindrance to growth. The general policy of limiting government scope makes sense with a caveat that we here we could not control for the size or composition of budgets.

One insight from these results is that growth researchers should use care in selecting indicators for measurement of economic well-being. The presence of non-linearities should also be tested, since we show that non-linearities are present in
the growth model in terms of interactions. Our results indicate that the somewhat contradictory results in the earlier literature, e.g. that institutions do not matter for growth or that Investment have a negative impact on growth in transition, might be due to an inappropriate choice of model. Even if we do not fully contradict these earlier results, we show that the effect is more complicated than previously argued.

Some issues remain, most strikingly the question of how best to measure human capital. In a limited scope of time and space, one might find, as we did, a set of countries where the quality of human capital is more or less constant. Due to this limited scope, however, the findings do not provide useful insights outside the scope of the sample. Clearly, current measures of human capital are far from perfect and better measurement methods are needed.
A Data

Table 4 lists the countries in our data set. The first 22 are not oil producers; the last three are.

<table>
<thead>
<tr>
<th>Non-oil</th>
<th>Hungary</th>
<th>Slovenia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albania</td>
<td>Kyrgyz Republic</td>
<td>Tajikistan</td>
</tr>
<tr>
<td>Armenia</td>
<td>Latvia</td>
<td>Turkmenistan</td>
</tr>
<tr>
<td>Belarus</td>
<td>Lithuania</td>
<td>Ukraine</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Moldova</td>
<td>Uzbekistan</td>
</tr>
<tr>
<td>Croatia</td>
<td>Poland</td>
<td>Oil</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>FYR Macedonia</td>
<td>Azerbaijan</td>
</tr>
<tr>
<td>Estonia</td>
<td>Romania</td>
<td>Kazakhstan</td>
</tr>
<tr>
<td>Georgia</td>
<td>Slovakia</td>
<td>Russia</td>
</tr>
</tbody>
</table>

Table 4: List of countries.

Table 5 lists the variables, their sources and their scales.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Scale</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>Real GDP per Worker</td>
<td>World Bank WDI</td>
</tr>
<tr>
<td>(Private) Investment</td>
<td>Share of the GDP</td>
<td>NBER PWT 6.2</td>
</tr>
<tr>
<td>General Government Expenditure</td>
<td>Share of the GDP</td>
<td>EBRD SEI</td>
</tr>
<tr>
<td>Human Development Indicator</td>
<td>From 0 to 1</td>
<td>UNDP</td>
</tr>
<tr>
<td>Transition Index</td>
<td>From 0 to 4+</td>
<td>EBRD Transition indicators</td>
</tr>
<tr>
<td>Economic Freedom</td>
<td>From 0-100</td>
<td>Quality of Gov. Institute</td>
</tr>
<tr>
<td>Fertility Rate</td>
<td>From 0-</td>
<td>UNDP</td>
</tr>
<tr>
<td>Gini Coefficients</td>
<td>From 0-1</td>
<td>Wider Institute</td>
</tr>
<tr>
<td>Upper Secondary Enrollment rates</td>
<td>From 0-100</td>
<td>World Bank WDI</td>
</tr>
</tbody>
</table>

Table 5: List of variables and their sources. The general government expenditure does not include public investments.

Table 6 gives the descriptive statistics on the variables that were used in the estimations.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Growth</th>
<th>Investment</th>
<th>Gov’t Cons.</th>
<th>EFI</th>
<th>EBRD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>std</td>
<td>mean</td>
<td>std</td>
<td>mean</td>
</tr>
<tr>
<td>Non-oil</td>
<td>mean</td>
<td>std</td>
<td>mean</td>
<td>std</td>
<td>mean</td>
</tr>
<tr>
<td>Albania</td>
<td>8.1</td>
<td>8.1</td>
<td>40.1</td>
<td>5.7</td>
<td>30.5</td>
</tr>
<tr>
<td>Armenia</td>
<td>11.1</td>
<td>5.5</td>
<td>10.9</td>
<td>1.0</td>
<td>23.3</td>
</tr>
<tr>
<td>Belarussia</td>
<td>6.0</td>
<td>4.4</td>
<td>15.4</td>
<td>2.0</td>
<td>30.5</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>3.1</td>
<td>5.7</td>
<td>7.7</td>
<td>2.7</td>
<td>38.2</td>
</tr>
<tr>
<td>Croatia</td>
<td>5.2</td>
<td>3.0</td>
<td>16.8</td>
<td>2.9</td>
<td>50.3</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>3.3</td>
<td>1.8</td>
<td>21.1</td>
<td>1.4</td>
<td>43.6</td>
</tr>
<tr>
<td>Estonia</td>
<td>7.5</td>
<td>2.2</td>
<td>15.8</td>
<td>1.9</td>
<td>36.5</td>
</tr>
<tr>
<td>Georgia</td>
<td>5.5</td>
<td>12.0</td>
<td>7.4</td>
<td>1.1</td>
<td>19.4</td>
</tr>
<tr>
<td>Hungary</td>
<td>3.6</td>
<td>1.2</td>
<td>20.8</td>
<td>2.3</td>
<td>49.9</td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td>2.3</td>
<td>4.2</td>
<td>8.8</td>
<td>1.4</td>
<td>30.4</td>
</tr>
<tr>
<td>Latvia</td>
<td>8.1</td>
<td>8.1</td>
<td>13.4</td>
<td>3.7</td>
<td>36.4</td>
</tr>
<tr>
<td>Lithuania</td>
<td>7.1</td>
<td>3.6</td>
<td>13.4</td>
<td>1.7</td>
<td>33.0</td>
</tr>
<tr>
<td>Moldova</td>
<td>4.2</td>
<td>6.8</td>
<td>8.7</td>
<td>0.5</td>
<td>38.3</td>
</tr>
<tr>
<td>Poland</td>
<td>4.9</td>
<td>2.0</td>
<td>19.8</td>
<td>2.1</td>
<td>44.9</td>
</tr>
<tr>
<td>FYR Macedonia</td>
<td>2.9</td>
<td>7.5</td>
<td>12.3</td>
<td>1.0</td>
<td>37.0</td>
</tr>
<tr>
<td>Romania</td>
<td>3.3</td>
<td>10.8</td>
<td>11.4</td>
<td>1.1</td>
<td>33.3</td>
</tr>
<tr>
<td>Slovakia</td>
<td>3.8</td>
<td>1.1</td>
<td>21.0</td>
<td>2.7</td>
<td>51.3</td>
</tr>
<tr>
<td>Slovenia</td>
<td>3.0</td>
<td>2.4</td>
<td>25.9</td>
<td>2.8</td>
<td>46.4</td>
</tr>
<tr>
<td>Tajikistan</td>
<td>3.2</td>
<td>6.2</td>
<td>6.2</td>
<td>2.0</td>
<td>18.8</td>
</tr>
<tr>
<td>Turkmenistan</td>
<td>5.5</td>
<td>13.2</td>
<td>24.5</td>
<td>4.8</td>
<td>20.6</td>
</tr>
<tr>
<td>Ukraine</td>
<td>2.5</td>
<td>8.6</td>
<td>8.9</td>
<td>1.5</td>
<td>38.4</td>
</tr>
<tr>
<td>Uzbekistan</td>
<td>1.8</td>
<td>2.5</td>
<td>6.0</td>
<td>0.9</td>
<td>35.1</td>
</tr>
<tr>
<td>Oil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>8.1</td>
<td>8.9</td>
<td>15.3</td>
<td>5.7</td>
<td>22.6</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>8.9</td>
<td>6.0</td>
<td>9.0</td>
<td>1.4</td>
<td>22.1</td>
</tr>
<tr>
<td>Russia</td>
<td>3.2</td>
<td>4.1</td>
<td>10.1</td>
<td>2.9</td>
<td>38.4</td>
</tr>
</tbody>
</table>

Table 6: Descriptive statistics
Figure 1: Human Development in 1995 and 2005 (own calculations).

Figure 2: Economic Freedom Index 1995 and 2005.
References


Durlauf, S.N., Johnson, P.A., 1995. Multiple regimes and cross-country behav-


# Convergence and Demographic Transition: A Club Approach*

November 12, 2009

## Contents

1. Introduction 118  
2. Data and Methods 120  
3. Demographic Clubs 124  
4. Convergence Within the Clubs 127  
   4.1 Unconditional $\beta$–Convergence 127  
   4.2 Cross-Section Dependence and Unit Root Tests 127  
5. Demographic Transition and Future Prospects 131  
6. Discussion 134  
   A Comparison of IPS and CIPS tests 139  
   B Countries and Clubs 140

---

*This paper has been co-authored by Professor Ulla Lehmijoki.
1 Introduction

The literature on the cross-country convergence of incomes is full of controversies and puzzles, but recently it has been suggested that the demographic transition may explain many conflicting observations (Galor 2007). The main argument is that differences in the timing of the demographic transition have segmented countries to the different regimes, and the simultaneous existence of these regimes makes the observations difficult to understand if their demographic background is not properly understood.

Demographics can increase our understanding in two ways. Since it is well known that convergence fails in heterogeneous samples of countries, it has been suggested earlier that countries should be classified into homogenous clubs according to common history (Baumol 1986), geographical location (Maddison 1994), mutual trade (Ben-David and Loewy 1998), or productivity thresholds (Azariadis and Drazen 1990). In this paper, we claim that in the post-war period countries were in different stages of their demographic transitions, so that classification according to these stages should generate homogenous clubs. The underlying tendency for convergence should then manifest itself within these clubs. On the other hand, since the demographic transition is a dynamic process, far from completed, many countries will move ahead toward more mature stages. Therefore, an important question is whether these movements imply a more equal distribution of world incomes, i.e., will there be convergence in the future?

In the demographic transition considerable changes in all demographic variables take place, but different authors emphasize different factors. Soares (2007), for example, maintains that infant mortality has the leading role: once it starts to fall fertility follows and the transition proceeds according to its internal laws. On the other hand, Ram (1998), Fogel (1994, 2004), Becker et al. (2005), and Lorenzen et al. (2008) claim that life lengthening is decisive as short-sighted and deterministic attitudes give way to optimistic views, whereas Bloom and Williamson (1998) stress the practical role of the dependence rate. In the “Unified Growth Theory", Galor and Weil (2000) emphasize the interplay between population growth and technical change. The number of demographic clubs has been disputed as well. Bloom and Williamson (1998) argue that this number should be limited to two, whereas Galor and Weil (2000) advocate three clubs, the “Malthusian", "Post Malthusian" and "Modern". Chesnais (1992), in turn, claims that the correct number of demographic stages is six.
In order to evaluate the role of the demographic transition in the convergence of incomes, one has to evaluate whether demographic clubs exist in the data, what their number is, and which demographic factors best discriminate between them. The next question concerns economic convergence within the clubs and, finally, one has to evaluate what the observations can tell us about the future.

To answer these questions, we identify the number and boundaries of the clubs by the regression tree method suggested by Breiman et al. (1984) and Durlauf and Johnson (1995). The main advantage is that this method chooses the variables and factors which most efficiently classify the data, thus providing important information about the clubs.

We apply various convergence tests to the clubs. The interest in growth and convergence has been extensive and a number of papers evaluate the topic from different perspectives. Li and Papell (1999) maintain that the failure to find convergence is due to structural changes, finding evidence for convergence if breaks are allowed. Their results are confirmed by Strazicich et al. (2004) for the OECD countries and by Cunado and Perez de Garcia (2006) for 43 African countries. Authors, such as Pedroni (2007) and Li and Papell (1999) find, however, that even within the OECD only a sub-set of countries is prone to convergence. While Pesaran (2007c) finds evidence for pair-wise convergence across the Middle East, North Africa and the Rest of the World, his findings, however, indicate that in a panel of countries one can rarely expect to find a convergence as a significant fraction of each sub-set fails to converge to the output mean. Hineline (2008) also highlights the merits of clustering the data when different countries obey different laws of motion.

Our results highlight the shortcomings of the cross-sectional tests and the potential problems in the first generation unit root tests when the data exhibits cross-section dependence, undermining the reliability of these tests. In the context of output convergence, this means that convergence may be found where it does not exist. We provide new evidence on the presence of cross-section dependence and show the variation of the convergence tests by different techniques from strong evidence in the traditional $\beta$-convergence tests to the failure to reject the non-convergence null by the second generation unit root tests. Thus, our results are best in line with the critical findings of Pedroni (2007), Li and Papell (1999), and Pesaran (2007c). To understand the implications of the demographic transition, we evaluate its progress in the future, predicting take-offs even among the poorest countries. These take-offs, however, will be insufficient to raise their incomes to meet the income of the rich, implying bi-polarization in the future.
The paper is structured as follows. Section 2 reviews the data and methods, Section 3 generates the clubs, Section 4 provides the convergence tests, and Section 5 discusses the future growth prospects. Section 6 discusses the merits of the findings.

2 Data and Methods

To investigate the role of demographics in the convergence of countries, we collect data for incomes and demographic variables. There are several candidates for demographic threshold variables, but we choose the total fertility rate, dependence rate, life expectancy, and infant mortality rate since their role is frequently discussed in the literature. Thus the variables are

\[ y_t = \log \text{ of real per capita GDP (international dollars, base year 2000).} \]

\[ GROWTH = \text{annual average growth rate of real per capita GDP defined by } \frac{(y_{t+T} - y_{t})}{T}. \]

\[ TFR = \text{Total fertility rate (children per woman).} \]

\[ DEP = \text{dependence rate (ratio of population aged 0-14 and 65+ per 100 population 15-64).} \]

\[ LIF = \text{Life expectancy (at birth, both sexes).} \]

\[ IMR = \text{Infant mortality (infant deaths per 1,000 live births).} \]

To keep the threshold variables exogenous, they are measured at the beginning of the research period, which extends from 1960 to 2003. Data for these variables is available for 85 countries. The countries that have experienced extreme economic or social changes are excluded. The demographic data come from the United Nations (2007) and the economic data from Heston et al. (2006).

To see whether multiple regimes should be taken seriously we make data splits according to the mean of each demographic threshold variable in country \( i \) \((x_i)\), and

\footnote{The excluded countries are the highest AIDS prevalence countries (Lesotho, South-Africa and Zimbabwe), the oil countries (OPEC members), and the East-European countries. We also exclude Rwanda and China because of the mass murders in the former and the population policy in the latter. The demographic data for Taiwan is replaced by that of South Korea which has quite a similar demographic history. The need to keep Taiwan arises because of the scarcity of countries with remarkable slow-downs in fertility.}
test whether \( GROWTH \) is identical in the sub-samples by estimating

\[
(y_{i,t_0+T} - y_{i,t_0}) / T = \varphi + \lambda x_i + \epsilon_i.
\]

(1)

The Wald-test for the similarity of the coefficients \( \varphi \) and \( \lambda \) in the sub-samples yields highly significant \( F \)-statistics for three of the four splits. Thus, we reject the similarity in favor of the sub-samples.

Even though the specification test suggests that some demographic clubs exist in the data, their number and boundaries are not adequately revealed by mechanical splits. Hence, clubs are discovered by using the regression tree analysis, suggested by Durlauf and Johnson (1995).\(^2\) This data-sorting method splits the range of the regressors to find the best piecewise linear model. In principle, the growth model might be different from one regime to another and this method helps to uncover the presence of non-linearities over space. We let the algorithm choose both the splitting variable and the split value (threshold) to generate the largest possible decreases in the model’s residual deviance. Only one-step look ahead and binary splits are used. Successive splits grow up a tree, starting from the root (the full sample) to the leaves (clubs).

Since we aim to minimize the sum of squared residuals of model (1), we need to set restrictions to this minimization problem. Otherwise the “optimal” number of clubs would probably be equal to \( N \). To choose the best number of clubs, several criteria are available. The \( V \)-fold validation method, for example, can be used to control for the potential over-fitting. This method, however, is computationally demanding since \( V \)-fold validation\(^3\) is done for all possible sizes of the tree, and its


\(^3\)V-fold cross validation is a technique for performing independent tree size tests. The data is split to samples (folds), their number of amounting \( V \) (typically 10). The samples should imitate the original data as closely as possible. Then a pseudo-data is created by leaving out one of the folds and a test tree is build. This tree is fitted to the remaining sample to assess its fit. The process is repeated 10 times for 10 different test samples. These ten trees give classification error for each tree size, according to which the reference tree can be pruned. It appears that this method requires sample size that is vastly greater than ours.
not available in the software we use to do this task. In our case we can resort to other criteria, as we also face the limits of the convergence tests, implying that the club size should not be too small. Therefore, we apply the pre-determined club-size criteria of ten members. A detailed description of the regression tree method is available in Breiman et al. (1980) and Durlauf and Johnson (1995).

Another important question is whether the generated clubs exhibit the convergence of incomes. A cross-section of countries is said to exhibit unconditional $\beta-$convergence if the estimated $\beta$ in the model

$$\text{Model 1: } (y_{i,t_0 + T} - y_{i,t_0}) / T = \alpha + \beta y_{i,t_0} + \varepsilon_{i,t}$$

is negative, indicating that economic growth in the poorer countries is faster than in the richer (Barro and Sala-i-Martin 1992). Evans (1998) first applied the panel unit root tests for the stationarity of output differences. This property can be tested by using three nested specifications from general to specific:

- **Model 2**: $\Delta (y_{i,t} - \tilde{y}_i) = \alpha_i + \theta_i t + \rho_i (y_{i,t-1} - \tilde{y}_{i-1}) + (u_{i,t} - \bar{u}_t)$,
- **Model 3**: $\Delta (y_{i,t} - \tilde{y}_i) = \alpha_i + \rho_i (y_{i,t-1} - \tilde{y}_{i-1}) + (u_{i,t} - \bar{u}_t)$,
- **Model 4**: $\Delta (y_{i,t} - \tilde{y}_i) = \rho_i (y_{i,t-1} - \tilde{y}_{i-1}) + (u_{i,t} - \bar{u}_t)$,

where $\tilde{y} = \frac{1}{N} \sum_{i=1}^{N} y_{i,t}$ and $u_{i,t}$ is iid. **Models 2 and 3** include a country-specific constant $\alpha_i$, necessary if some slowly-changing factor wedges the incomes from the mean. **Model 2** also includes a country-specific trend $\theta_i t$, addressing time-related factors, such as the diffusion of technology, which may take place at different pace in different countries (Lee et al. 1997). It is often necessary to allow this kind of heterogeneity even within clubs since a complete control of heterogeneity by clustering may not be possible. Only the test with no intercept and trend (**Model 4**) always refers to decreasing income gaps, i.e., to unconditional convergence, whereas **Models 2 and 3** only refer to the conditional one. For discussion see Pesaran (2007b) and Pedroni (2007).

In **Models 2-4**, country $i$ converges to the mean (has a stationary time series of income differences) if the estimate for $\rho_i$ is negative, but several test variants exist...
in terms of similarity of this estimate across countries. Levin, Lin and Chu (LLC, 2002) assume convergence at a common rate, i.e., $\rho_i = \rho$ for all $i$. Im, Pesaran and Shin (IPS, 2003) propose a test statistics which builds on the Augmented Dickey Fuller (ADF) test. This test, as well as the Fisher inverse square test by Maddala and Wu (M&W, 1999) and Choi’s (2001) inverse normal test all assume individual unit root processes, indicating that countries may converge at different rates and some countries may not converge at all. For the convergence of the sample it is then enough to show that $\tilde{\rho} < 0$. The difference between LLC and the other tests is that LLC pools the data while the other tests pool the test statistics, hence different assumptions about $\rho_i$.

Recently, Pesaran (2007a) has criticized the use of these so called first generation panel unit root tests because they do not account for cross-section dependence, arising across countries due to spatial and spill over effects or unobserved common factors (Baltagi and Pesaran 2007). Although IPS and Choi both allow for a limited amount of cross-section dependence due to demeaning in the presence of common time effects (common business cycles, for example), demeaning does not help if reaction to shocks differs across countries.\footnote{Assume that $u_{i,t}=\lambda_j f_t + \varepsilon_{i,t}$. If $\tilde{\lambda}_j = \lambda$ for all $j$ then demeaning, as suggested in Models 2 to 4, is enough to whiten the error term, otherwise we need to resort to second generation unit root tests.} Pesaran (2007a) investigates the properties of the IPS, M&W, and Choi tests in the presence of cross-section dependence by Monte Carlo simulations. With low dependence, IPS and Choi perform reasonably well, whereas M&W begins to work when $T$ increases. With high cross-section dependence, all tests tend to over-reject the no-convergence null.

To correct for the bias rising from cross-section dependence, Pesaran (2007a) proposes a modified IPS in the presence of a single unobserved common factor, but this does not come without costs: if no dependence exists, the corrected IPS (CIPS) performs worse than the original test.\footnote{The Augmented Dickey-Fuller (of order one) in the presence of individual effects can be written as $\Delta z_{i,t} = a_i + \phi_i z_{i,t-1} + c_{i1} \Delta z_{i,t-1} + \varepsilon_{i,t}$. This regression gives the base for the IPS-tests. The modified ADF-test in the presence of single common factor $f_t$ can be estimated from $\Delta z_{i,t} = a_i + \phi_i z_{i,t-1} + b_i \tilde{z}_{t-1} + c_{i1} \Delta \tilde{z}_t + \varepsilon_{i,t}$. While the standard ADF regression uses the lagged} While the CIPS test rarely over-rejects the
null, its power is often relatively low, i.e., if the null hypothesis is false, the test may fail to reject it. The second pitfall is that since Pesaran’s test builds on IPS, the unconditional convergence in Model 4 can not be tested with it.

To discover the cross-section dependence on the data, the CD-test proposed by Pesaran (2007a) calculates

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left( \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\gamma}_{ij} \right), \quad (2)$$

where $T$ and $N$ are the number of observations in time and cross-sections, and $\hat{\gamma}_{ij}$ is the residual correlation between countries $i$ and $j$, these residuals being obtained from individual ADF(p) regression$^6$ The statistics of this test is normally distributed with $N(0, 1)$, but the drawback is that it lacks power if the population average pairwise correlation is (close to) zero. Another test, proposed by Breusch and Pagan (1980)

$$LM = T \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\gamma}_{ij}^2 \quad (3)$$

is based on $\chi^2_{N(N-1)/2}$ distribution. While this test is not affected by the zero averages, it is likely to exhibit substantial size distortions when $N$ is large and $T$ is small.

3 Demographic Clubs

Figure 1 reports the results of the regression tree analysis. The first split is in terms of life expectancy, the split value being $LIF = 48.01$. Countries with $LIF_i < \overline{\text{difference(s) of } z_{i,t}}$, the modified test uses $\overline{z}_{t-1}$ and $\Delta \overline{z}_t$, i.e. the cross-section averages, as a proxy for the common factor $f_t$ (for more details see Pesaran 2007a and our Appendix A). Then for the (C)IPS unit root test one calculates the mean of the individual specific $t$-statistics.

$^6$Pesaran et al. (2007), who proposed the use of a bias-adjusted LM test, show that this test is more robust than the other two. The problem with the adjusted LM test is that it assumes strong exogeneity, which does not hold in our data.
48.01 constitute the first club (35 countries).\(^7\) The second split is in terms of life expectancy again, the split value being \(LIF = 55.785\), thus partitioning the rest of the sample into the sub-samples of 12 and 38 countries \((LIF_i < 55.785\) and \(LIF_i \geq 55.785\)). The next split violated the club-size requirement, whereby the (maximum) number of clubs is three. Most developing countries are allocated to Club I, the East-Asian Tigers to Club II, and the Western countries to Club III. A complete list of countries and clubs with demographic and economic statistics is given in Appendix A.

Figure 1: The regression tree. The left arrow indicates the observations for which the split variable < Split Value. The right arrow indicates the observations for which the split variable \(\geq\) Split Value.

The order by which the threshold variables enter to partition the data indicates their importance in terms of the dependent variable. Our study emphasizes the

\(^7\)The algorithm then suggests the second split in terms of the dependence rate with the split value \(DEP = 94.94\) but, unfortunately, the number of members in the club with \(DEP_i \geq 94.94\) is too low (eight countries), thus violating the pre-determined minimum size for the clubs. To be able to continue the partition further we exclude the dependence rate from the threshold candidates.
importance of life expectancy. Two types of intuitive explanation arise. The first
derives from the role of the physical capacity of workers (Fogel 1994, 2004) and
highlights better incentives to invest in human capital associated with the lengthen-
(2009) argue, however, that the causal link from life expectancy to economic growth
may be weak as the rise in the supply of a qualified labor force may decrease its
marginal productivity, in particular if investments are weak. A more reliable ex-
planation is thus that because the diffusion of production and health technologies
go hand in hand, life expectancy acts as an indicator for adaption of latest tech-
nologies to increase income (Soares 2007). This explanation also sheds light on the
growth differentials between the clubs. Given that technological diffusion depends
on the gap to the leaders and the ability to adopt information, Club II, where both
conditions hold, exhibits highest growth (Cf. Figure 1), while Club I, with the
largest gap, suffers from deficient adoption. Club III is at the frontier. Naturally,
there might be gaps within the clubs leading to convergence as recently illustrated
in Lehmioki (2009).

The derived number of the demographic clubs is identical with that suggested by
Galor and Weil (2000). As the discussion above suggests, their driving variables, hu-
man capital and technical change, may not be so different from our life expectancy.
Furthermore, the increasing life-expectancy status from Club I to Club III seems to
confirm the interpretation of our clubs as successive demographic stages, denoted
as “Malthusian”, “Post Malthusian” and “Modern” by Galor and Weil (2000).8

8Fiaschi and Lavezzi (2007) also find three clubs out of 122 countries, although based on initial
incomes. The number of the literacy and income-based regimes (96 countries) in Durlauf and
Johnson (1995) is four, their club 1 being a sub-set of our Club I, while their club 4 is a sub-set
of our Club III. Their intermediate clubs consist of countries from all of our clubs.
4 Convergence Within the Clubs

4.1 Unconditional $\beta$–Convergence

In this section we concentrate on the convergence of incomes within demographic clubs, starting from the unconditional $\beta$-convergence defined in Model 1. Table 1 shows that the estimated $\beta$ in the full sample is positive and significant, i.e., the sample diverges. By contrast, $\beta$ is negative in clubs although significant at the 5% level only in Clubs I and III. Convergence in samples similar to Club III has been shown in earlier studies but the now-shown $\beta$-convergence in Club I is a new result. The estimated $\beta$ in Club II is large but significant only at the 8% level. A closer examination shows that its members have changed their relative positions and the correlation of rank orders in 1960 and 2003 is -0.06, i.e., there is “leapfrogging” in this club. This possibility was discussed by Sala-i-Martin (1996) and is now exemplified in Club II.

<table>
<thead>
<tr>
<th>club</th>
<th>Model 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
</tr>
<tr>
<td>Full sample</td>
<td>0.354</td>
</tr>
<tr>
<td>Club I</td>
<td>-0.640</td>
</tr>
<tr>
<td>Club II</td>
<td>-1.912</td>
</tr>
<tr>
<td>Club III</td>
<td>-0.479</td>
</tr>
</tbody>
</table>

Table 1: The results for Model 1

4.2 Cross-Section Dependence and Unit Root Tests

To evaluate the presence of cross-sectional dependence, we ran the ADF(p) regressions with the individual values of $p$ to compute the pair-wise cross-section correlations for the residuals (Pesaran 2007a). Table 2 shows that the average correlations are -0.021, -0.061, and -0.01 for Clubs I – III respectively.\(^9\) The CD statistics for

\(^9\)Since we are testing the convergence hypothesis, we ran the ADF(p) regressions for the demeaned data. If the cross-section dependence is still present, it is off more complex form than expected by the first generation tests so that they are biased.
The first two values are highly significant, while the value for the third is not significant at 5% level (CD < |1.96|). The CD test indicates strong cross-section dependence in Clubs I and II. Moreover, since the power of this test is low when the average correlations tend to zero as in Club III, it may be the case that Club III is prone to cross-section dependence as well. The LM test statistics are also high and significant, although Clubs I and III are probably too large to yield reliable results. Hence, there is some cross-section dependence on the data, rendering the use of second generation tests.

Table 3 summarizes the first generation unit root tests for Model 2, showing that no convergence exists in the full sample. The tests also accept the null of no convergence for Clubs II and III, while the LLC and M&W tests offer weak support for convergence in Club I. The second generation test (Pesaran’s CIPS), on the other hand, suggests that the null of non-stationarity cannot be rejected for any of the clubs. The power of the CIPS test is relatively low in the presence of trends, so that the rejection of non-stationarity is somewhat expected. This rejection, however, is supported by the conventional IPS and the Choi tests, which are rather robust in the presence of moderate cross-section dependence. Moreover, as they tend to over-reject the null in this case, the fact that they do not strengthen the finding.

---

10 All first generation tests were performed by Eviews and cross-checked with Stata. We allowed automatic selection of lags based on Schwartz Information Criteria and when kernels had to be used, we used Newey-West bandwidth selection by Bartlett. The second generation test is only available in Stata; individual lags were allowed here.

11 The critical values of CIPS-test are tabulated by Pesaran (2007a) Tables II(a)-II(c).
Table 3: Results for Model 2.

Table 4 reports the results with intercept but without trend (Model 3), i.e. the results for conditional convergence. The first generation unit root tests (with the exception of IPS) suggest that Club III exhibits convergence\(^{12}\) but the second generation test (Pesaran’s CIPS) is far from being statistically significant. The first generation unit root tests that are most robust in the presence of moderate cross-section dependence (IPS and Choi) also fail to indicate convergence. LLC and Choi also suggest convergence for the data as a whole, but this finding is not supported by the more robust tests.

Table 4: Results for Model 3.

Table 5 presents the results for unconditional convergence (Model 4) replicating the common finding that the rich countries exhibit convergence while convergence is not found in other clubs. Unfortunately, this test cannot be performed by IPS so no second generation test results are yet available.

\(^{12}\)Choi also suggest convergence for the data as a whole.
To summarize, we find unconditional $\beta$-convergence ($Model 1$) in all clubs, indicating that the regression tree analysis has succeeded in uncovering three traditional convergence clubs on the demographic basis alone. First generation unit root tests which provide opportunities to control for the heterogeneity within the clubs ($Models 2$ and $3$) also support conditional convergence in Club I but none in Club II. The tests, however, are likely to be biased by the cross-section dependence and the more robust tests do not find support for conditional convergence in these clubs. Furthermore, no support for unconditional convergence ($Model 4$) can be found for these clubs. Maybe the most interesting is Club III as the first generation tests give some support for conditional convergence ($Model 3$), this result being undermined by the second generation test again. On the other hand, the support for unconditional convergence is strong ($Model 4$). Unfortunately, the unavailability of the second generation test for $Model 4$ leaves this question open as cross-section dependence is also indicated in this club.

How do we now conclude with the argument that convergence, if it exists, should manifest itself in homogenous clubs? From the technical point of view, our findings are in line with the development of the literature, running from the earlier findings of no $\beta$-convergence in broad samples but some in clubs (Baumol 1986, Mankiw et al. 1992, Maddison 1994, Ben-David and Loewy 1998) to the markedly skeptical findings from the recent tests (Li and Papell 1999, Pedroni 2007, Pesaran 2007c), indicating that the quest on convergence is far from settled.\footnote{To compare our results with earlier works, we performed the unit root tests on the club 4 of...} On the other hand, a
demographic explanation for the mixed evidence can also be found since the demo- 
graphic transition is a continuous rather than discrete process, necessarily leaving 
some timing-heterogeneity in the clubs. Chesnais (1992) has also emphasized the 
varying phases of the transition across countries, implying that by controlling for 
the initial states alone we may have failed to control for the differences in the phase. 
Such control is a necessary future task, although challenging, as the phase of the 
transition may be dictated by economic growth itself.

5 Demographic Transition and Future Prospects

In spite of the mixed evidence within the clubs, the demographic approach helps us 
to evaluate the future growth differentials between the clubs. To illustrate, consider 
the club-specific average incomes at the beginning and end of the period 1960-2000 
and compare Clubs I and III, for example. Because of the considerable differences 
in the growth rates (0.74% versus 2.49%), the income gap increased from five to 
tenfold, this increase being exacerbated by a massive demographic expansion in 
the former club (Figure 2). By contrast, the income gap between Clubs III and II 
decreased from four to 2.5 fold due to rapid growth (3.21%) in the latter club.

The theory of the demographic transition presupposes that countries move ahead 
to more mature stages, i.e., the demographic clubs are transitional rather than 
permanent, and multiple steady states may not be present in the data (Galor and 
Weil 2000, Galor 2007). Unfortunately, none of the techniques above can identify 
the data generating processes, but the dynamics of the demographic variables can 
give some hints. Figure 3 shows that life expectancy has increased everywhere such 
that its average value in Club I (Club II) in 2003 exceeded that in Club II (Club 
III) in 1960. Analogous information is given by total fertility, infant mortality, and

Durlauf and Johnson (1995), as it is a sub-set of our Club III. While they find that club 4 exhibited 
unconditional $\beta$-convergence at 1960-1985, we find that it also shows both $\beta$- convergence and the 
convergence in the first generation unit root tests but no convergence in the second generation 
the by the dependence rate, all of which have greatly decreased since 1960 as most countries have reached the higher stages, thus supporting transitional clubs.

This makes some future explorations possible. Consider a new classification derived by applying the earlier boundaries of the clubs to the values of the life expectancies in 2003. This classification shows that only six countries still stay in Club I and twelve in Club II, while all other countries (67) have proceeded to Club III.\(^\ast\) One can now predict the average incomes, say, in 2040 illustrated in Figure 4.\(^\ast\) A comparison of Figures 4 and 2 shows that the future does not replicate the past. On the contrary, while the average income in Club II approached that of Club III in the 1960-2003 period, in the 2004-2040 period it will fall behind since the inherited low incomes provide a limited basis regardless of rapid growth (3.21%). Thus, the countries which have migrated from Club I to Club II will experience a

\(^\ast\)The countries staying in Club I are Burundi, Cote d’Ivoire, Guinea-Bissau, Mozambique, Malawi, and Zambia. The countries in Club II are Benin, Burkina Faso, Cameroon, Congo, Ethiopia, Guinea, Kenya, Mali, Niger, Chad, Tanzania, and Uganda.

\(^\ast\)There are few attempts to predict the evolution of the future incomes, and to our knowledge, none in a cross-country set-up. Holz (2008), however, uses Chinese real GDP growth rates from 1978-2000 to extrapolate the future to evaluate when the size of Chinese economy would surpass that of US in absolute terms.
take-off, but this take-off is only in terms of the country’s own history and does not raise its income closer to the more advanced countries which have already proceeded too far. Therefore, the catch-up opportunity gained by the countries that arrived in Club II in the post-war period is not open to those who will arrive later. The window of opportunity has been closed.

Figure 4: The per capita GDP in 2003 (grey bar) and 2040 (black bar).

Club III may also experience changes as 29 of its 67 members will be new-comers.
While these new-comers may grow at the rate which was typical of this club in the past, the old members may meet new growth-hampering problems, such as ageing, not visible in the 1960-2003 period yet. Hence, new convergence tendencies may arise within Club III. This, together with the limited perspectives from Clubs I and II, indicates bi-polarization rather than convergence of world incomes as a whole. The number of countries in Clubs I and II, however, has decreased and their proportion of population will be smaller than before.

6 Discussion

This paper explores the role of demographics in the post-war growth and convergence between countries. Differences in the timing of the demographic transition have segmented countries into the different regimes or clubs and the simultaneous existence of these clubs makes the concept of convergence meaningless if the existence of such clubs neglected. We evaluate the relevance of demographics by classifying countries into demographic clubs by the regression tree method. The discriminating variable turns out to be life expectancy, probably because of its role as an indicator of technical diffusion and as a necessary condition for investment in human capital, classifying countries into three demographic clubs.

The traditional \( \beta \)-convergence and first generation unit root tests offer support for convergence in demographic clubs but the second generation unit root test undermines this result so that the evidence is equivocal. This finding is in line with the trends in the literature, where the earlier tests suggest convergence but the most recent tests show that homogenous clubs are difficult to identify even among the OECD countries. The rapid progress in the field of panel estimation techniques may shed further light on this subject in the future. On the other hand, the demographic transition as a source of mixed evidence calls for further research efforts because its phase may differ across countries and controlling for the initial state alone may not lead to convergence clubs. Hence, a control for the heterogeneity from the different phases may be necessary, although challenging, as the phase of transition may be dictated by economic growth itself.
Continuous demographic transition elevates countries to higher clubs, providing important implications for future incomes. Unfortunately, this information gives no unequivocal support to convergence of world incomes as the income gaps have already widened to such an extent that even the take-offs, typical to the second stage or club, is unable to arise the incomes of the poor sufficiently. Thus, new economic miracles will hardly arise on a demographic basis alone. Therefore, economic policies should be targeted to help the countries which remain in the lowest clubs. The good news from the analysis here is that the ever richer majority of countries has better opportunity to manage this task since the number of these countries is small and their population share is low.
References


A Comparison of IPS and CIPS tests

In the presence of fixed effects, the augmented Dickey-Fuller test of order one is based on

$$\Delta z_{it} = a_i + \phi_i z_{i,t-1} + c_{i1} \Delta z_{i,t-1} + \varepsilon_{it}$$

(4)

where the lagged difference accounts for the short-run dynamics of the errors. Then for the IPS unit root test we calculate mean of the individual specific t-statistics

$$\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_i$$

where

$$t_i = \frac{\Delta z'_i M_e z_{i,-1}}{\tilde{\sigma}_i (\tilde{z}'_{i,-1} M_e \tilde{z}_{i,-1})^{1/2}}$$

is the Dickey-Fuller t-statistic of cross section unit $i$, where $M_e = I_T - \bar{W}' (\bar{W}' \bar{W})^{-1} \bar{W}'$, \( \bar{W} = (\tau, \Delta \tilde{z}, \tilde{z}, \tilde{z}_{-1}), \tau = (1, 1, ..., 1)', \Delta \tilde{z} = (\Delta \tilde{z}_1, \Delta \tilde{z}_2, ..., \Delta \tilde{z}_T)', \tilde{z}_{-1} = (\tilde{z}_0, \tilde{z}_1, ..., \tilde{z}_{T-1})' \)

and $\tilde{\sigma}_i^2 = \frac{\Delta z'_i M_e \Delta z_i}{T-4}$.

Now assume that the true data generating process is

$$z_{it} = (1 - \psi_i) \mu_i + \psi_i z_{i,t-1} + u_{it},$$

(5)

where the error term has the single-factor structure

$$u_{it} = \gamma_i f_t + \varepsilon_{it}.\quad (6)$$

We may now re-write (5) and (6) as

$$\Delta z_{it} = a_i + \phi_i z_{i,t-1} + \gamma_i f_t + \varepsilon_{it},$$

where $a_i = (1 - \psi_i)$, $\phi_i = \psi_i$ and $\Delta z_{it} = z_{it} - z_{i,t-1}$. Then if $\bar{\gamma} = \frac{1}{N} \sum_{i=1}^{N} \gamma_i \neq 0$, according to Pesaran (2007a), the common factor $f_t$ can be proxied by $\tilde{z}$ and its lagged values for $N$ sufficiently large. The unit root test is based on the following regression

$$\Delta z_{it} = a_i + \phi_i z_{i,t-1} + b_i \tilde{z}_{t-1} + c_{i1} \Delta \tilde{z}_t + \varepsilon_{it}.\quad (7)$$

Comparing (4) and (7) shows that while the former is augmented by the lagged difference of $z_t$ (i.e. $\Delta z_{i,t-1}$) the latter is augmented by the lagged mean of $z_t$ (i.e. $\tilde{z}_{t-1}$) and the difference of $\tilde{z}_t$ (i.e. $\Delta \tilde{z}_t$).
## B Countries and Clubs

<table>
<thead>
<tr>
<th>Country</th>
<th>GDP60</th>
<th>LIF60</th>
<th>IMR60</th>
<th>TFR60</th>
<th>DEP60</th>
<th>GDP03</th>
<th>LIF03</th>
<th>IMR03</th>
<th>TFR03</th>
<th>DEP05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benin</td>
<td>956</td>
<td>41.48</td>
<td>165</td>
<td>6.96</td>
<td>99.49</td>
<td>1345</td>
<td>54.68</td>
<td>105</td>
<td>5.82</td>
<td>88.32</td>
</tr>
<tr>
<td>Bolivia</td>
<td>2431</td>
<td>42.65</td>
<td>164</td>
<td>6.63</td>
<td>85.35</td>
<td>3006</td>
<td>64.06</td>
<td>55</td>
<td>3.91</td>
<td>74.30</td>
</tr>
<tr>
<td>Burkina F.</td>
<td>768</td>
<td>36.36</td>
<td>197</td>
<td>7.00</td>
<td>76.47</td>
<td>1071</td>
<td>50.81</td>
<td>109</td>
<td>6.33</td>
<td>97.05</td>
</tr>
<tr>
<td>Burundi</td>
<td>677</td>
<td>41.29</td>
<td>149</td>
<td>6.80</td>
<td>88.85</td>
<td>763</td>
<td>47.56</td>
<td>106</td>
<td>6.81</td>
<td>91.35</td>
</tr>
<tr>
<td>Cameroon</td>
<td>1947</td>
<td>41.52</td>
<td>145</td>
<td>5.90</td>
<td>78.97</td>
<td>2713</td>
<td>49.86</td>
<td>90</td>
<td>4.84</td>
<td>82.82</td>
</tr>
<tr>
<td>Chad</td>
<td>1142</td>
<td>40.79</td>
<td>169</td>
<td>6.09</td>
<td>79.30</td>
<td>884</td>
<td>50.50</td>
<td>124</td>
<td>6.51</td>
<td>96.80</td>
</tr>
<tr>
<td>Comoros</td>
<td>1354</td>
<td>43.44</td>
<td>154</td>
<td>6.91</td>
<td>90.47</td>
<td>1278</td>
<td>63.21</td>
<td>57</td>
<td>4.83</td>
<td>80.65</td>
</tr>
<tr>
<td>Congo</td>
<td>1010</td>
<td>46.69</td>
<td>130</td>
<td>5.99</td>
<td>81.50</td>
<td>1420</td>
<td>53.15</td>
<td>75</td>
<td>4.76</td>
<td>82.22</td>
</tr>
<tr>
<td>Cote d’Iv.</td>
<td>1334</td>
<td>43.67</td>
<td>154</td>
<td>7.31</td>
<td>85.97</td>
<td>2019</td>
<td>46.90</td>
<td>121</td>
<td>5.00</td>
<td>81.47</td>
</tr>
<tr>
<td>Egypt</td>
<td>1469</td>
<td>45.74</td>
<td>180</td>
<td>7.07</td>
<td>84.30</td>
<td>4759</td>
<td>70.00</td>
<td>35</td>
<td>3.14</td>
<td>61.72</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>400</td>
<td>40.06</td>
<td>160</td>
<td>6.90</td>
<td>89.08</td>
<td>688</td>
<td>50.96</td>
<td>95</td>
<td>5.73</td>
<td>89.93</td>
</tr>
<tr>
<td>Gambia</td>
<td>722</td>
<td>32.99</td>
<td>223</td>
<td>6.50</td>
<td>74.64</td>
<td>937</td>
<td>58.18</td>
<td>79</td>
<td>5.12</td>
<td>81.42</td>
</tr>
<tr>
<td>Ghana</td>
<td>412</td>
<td>45.62</td>
<td>124</td>
<td>6.90</td>
<td>88.65</td>
<td>1440</td>
<td>58.64</td>
<td>63</td>
<td>4.32</td>
<td>74.34</td>
</tr>
<tr>
<td>Guatemala</td>
<td>2494</td>
<td>45.47</td>
<td>127</td>
<td>6.85</td>
<td>93.87</td>
<td>3805</td>
<td>69.11</td>
<td>38</td>
<td>4.55</td>
<td>90.27</td>
</tr>
<tr>
<td>Guinea</td>
<td>3072</td>
<td>35.12</td>
<td>201</td>
<td>7.00</td>
<td>83.19</td>
<td>2887</td>
<td>53.92</td>
<td>112</td>
<td>5.81</td>
<td>86.72</td>
</tr>
<tr>
<td>Guinea-B.</td>
<td>493</td>
<td>33.99</td>
<td>196</td>
<td>5.99</td>
<td>70.97</td>
<td>584</td>
<td>45.56</td>
<td>120</td>
<td>7.11</td>
<td>101.88</td>
</tr>
<tr>
<td>Honduras</td>
<td>1715</td>
<td>46.34</td>
<td>135</td>
<td>7.42</td>
<td>97.14</td>
<td>2291</td>
<td>68.77</td>
<td>31</td>
<td>3.68</td>
<td>79.02</td>
</tr>
<tr>
<td>India</td>
<td>892</td>
<td>41.87</td>
<td>140</td>
<td>5.81</td>
<td>76.75</td>
<td>2990</td>
<td>63.07</td>
<td>62</td>
<td>3.08</td>
<td>61.18</td>
</tr>
<tr>
<td>Jordan</td>
<td>4151</td>
<td>46.82</td>
<td>110</td>
<td>8.00</td>
<td>94.17</td>
<td>3743</td>
<td>71.41</td>
<td>23</td>
<td>3.49</td>
<td>67.75</td>
</tr>
<tr>
<td>Kenya</td>
<td>1179</td>
<td>46.29</td>
<td>117</td>
<td>8.12</td>
<td>100.78</td>
<td>1218</td>
<td>51.32</td>
<td>69</td>
<td>5.00</td>
<td>82.90</td>
</tr>
<tr>
<td>Madagascar</td>
<td>1268</td>
<td>39.87</td>
<td>155</td>
<td>6.80</td>
<td>87.14</td>
<td>759</td>
<td>57.54</td>
<td>74</td>
<td>5.24</td>
<td>88.51</td>
</tr>
<tr>
<td>Malawi</td>
<td>460</td>
<td>37.80</td>
<td>203</td>
<td>7.00</td>
<td>93.44</td>
<td>771</td>
<td>45.47</td>
<td>100</td>
<td>5.99</td>
<td>100.30</td>
</tr>
<tr>
<td>Mali</td>
<td>797</td>
<td>36.13</td>
<td>216</td>
<td>7.11</td>
<td>80.65</td>
<td>1184</td>
<td>52.11</td>
<td>137</td>
<td>6.71</td>
<td>105.49</td>
</tr>
<tr>
<td>Morocco</td>
<td>1299</td>
<td>46.68</td>
<td>145</td>
<td>7.15</td>
<td>90.14</td>
<td>4061</td>
<td>69.79</td>
<td>37</td>
<td>2.50</td>
<td>55.06</td>
</tr>
<tr>
<td>Mozambique</td>
<td>838</td>
<td>35.01</td>
<td>185</td>
<td>6.50</td>
<td>82.05</td>
<td>1452</td>
<td>43.79</td>
<td>106</td>
<td>5.49</td>
<td>90.19</td>
</tr>
<tr>
<td>Nepal</td>
<td>800</td>
<td>38.50</td>
<td>187</td>
<td>6.06</td>
<td>79.22</td>
<td>1441</td>
<td>61.66</td>
<td>63</td>
<td>3.64</td>
<td>74.27</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>4428</td>
<td>47.02</td>
<td>131</td>
<td>7.33</td>
<td>100.95</td>
<td>3409</td>
<td>71.07</td>
<td>26</td>
<td>2.97</td>
<td>71.96</td>
</tr>
<tr>
<td>Niger</td>
<td>1167</td>
<td>39.08</td>
<td>196</td>
<td>7.90</td>
<td>100.83</td>
<td>834</td>
<td>54.77</td>
<td>118</td>
<td>7.43</td>
<td>104.51</td>
</tr>
<tr>
<td>Pakistan</td>
<td>801</td>
<td>46.61</td>
<td>148</td>
<td>6.28</td>
<td>82.70</td>
<td>2593</td>
<td>63.78</td>
<td>75</td>
<td>3.95</td>
<td>69.64</td>
</tr>
<tr>
<td>Peru</td>
<td>3129</td>
<td>47.68</td>
<td>136</td>
<td>6.85</td>
<td>87.80</td>
<td>4351</td>
<td>70.09</td>
<td>29</td>
<td>2.67</td>
<td>59.59</td>
</tr>
<tr>
<td>Country</td>
<td>GDP60</td>
<td>LIF60</td>
<td>IMR60</td>
<td>TFR60</td>
<td>DEP60</td>
<td>GDP03</td>
<td>LIF03</td>
<td>IMR03</td>
<td>TFR03</td>
<td>DEP05</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Senegal</td>
<td>1776</td>
<td>40.39</td>
<td>186</td>
<td>7.00</td>
<td>82.23</td>
<td>1407</td>
<td>61.78</td>
<td>69</td>
<td>5.17</td>
<td>86.70</td>
</tr>
<tr>
<td>Tanzania</td>
<td>502</td>
<td>43.63</td>
<td>136</td>
<td>6.80</td>
<td>93.05</td>
<td>912</td>
<td>49.90</td>
<td>78</td>
<td>5.60</td>
<td>89.99</td>
</tr>
<tr>
<td>Togo</td>
<td>833</td>
<td>42.55</td>
<td>159</td>
<td>7.10</td>
<td>88.96</td>
<td>789</td>
<td>57.64</td>
<td>93</td>
<td>5.31</td>
<td>86.37</td>
</tr>
<tr>
<td>Uganda</td>
<td>873</td>
<td>43.95</td>
<td>130</td>
<td>6.90</td>
<td>94.27</td>
<td>1113</td>
<td>48.01</td>
<td>83</td>
<td>6.73</td>
<td>107.67</td>
</tr>
<tr>
<td>Zambia</td>
<td>910</td>
<td>45.07</td>
<td>127</td>
<td>6.62</td>
<td>90.17</td>
<td>946</td>
<td>39.44</td>
<td>101</td>
<td>5.60</td>
<td>94.69</td>
</tr>
<tr>
<td>Club1</td>
<td>1386</td>
<td>41.95</td>
<td>159</td>
<td>6.84</td>
<td>87.24</td>
<td>1882</td>
<td>56.81</td>
<td>79</td>
<td>4.99</td>
<td>83.92</td>
</tr>
<tr>
<td>Brazil</td>
<td>2644</td>
<td>54.51</td>
<td>109</td>
<td>6.15</td>
<td>87.02</td>
<td>7205</td>
<td>71.18</td>
<td>27</td>
<td>2.33</td>
<td>51.43</td>
</tr>
<tr>
<td>Cape Verde</td>
<td>1417</td>
<td>51.98</td>
<td>105</td>
<td>7.00</td>
<td>91.67</td>
<td>5117</td>
<td>70.38</td>
<td>29</td>
<td>3.71</td>
<td>78.03</td>
</tr>
<tr>
<td>Dominican R.</td>
<td>2080</td>
<td>51.69</td>
<td>124</td>
<td>7.32</td>
<td>101.67</td>
<td>6904</td>
<td>71.03</td>
<td>34</td>
<td>2.94</td>
<td>64.08</td>
</tr>
<tr>
<td>El Salvador</td>
<td>2991</td>
<td>50.42</td>
<td>123</td>
<td>6.85</td>
<td>94.99</td>
<td>4751</td>
<td>70.90</td>
<td>26</td>
<td>2.86</td>
<td>65.55</td>
</tr>
<tr>
<td>Korea</td>
<td>1458</td>
<td>54.20</td>
<td>70</td>
<td>5.63</td>
<td>82.72</td>
<td>17597</td>
<td>77.23</td>
<td>5</td>
<td>1.24</td>
<td>39.05</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1801</td>
<td>53.92</td>
<td>63</td>
<td>6.72</td>
<td>94.89</td>
<td>12133</td>
<td>73.18</td>
<td>10</td>
<td>2.84</td>
<td>55.63</td>
</tr>
<tr>
<td>Philippines</td>
<td>2039</td>
<td>53.20</td>
<td>96</td>
<td>6.85</td>
<td>95.64</td>
<td>3575</td>
<td>70.46</td>
<td>27</td>
<td>3.51</td>
<td>66.61</td>
</tr>
<tr>
<td>Syria</td>
<td>837</td>
<td>49.22</td>
<td>119</td>
<td>7.60</td>
<td>98.61</td>
<td>2016</td>
<td>73.19</td>
<td>18</td>
<td>3.42</td>
<td>65.99</td>
</tr>
<tr>
<td>Taiwan</td>
<td>1444</td>
<td>54.20</td>
<td>70</td>
<td>5.63</td>
<td>82.72</td>
<td>19885</td>
<td>77.23</td>
<td>5</td>
<td>1.24</td>
<td>39.05</td>
</tr>
<tr>
<td>Thailand</td>
<td>1059</td>
<td>54.89</td>
<td>83</td>
<td>6.40</td>
<td>87.39</td>
<td>7274</td>
<td>68.80</td>
<td>12</td>
<td>1.83</td>
<td>41.79</td>
</tr>
<tr>
<td>Tunisia</td>
<td>2103</td>
<td>48.34</td>
<td>155</td>
<td>7.25</td>
<td>90.59</td>
<td>7601</td>
<td>73.16</td>
<td>22</td>
<td>2.01</td>
<td>47.72</td>
</tr>
<tr>
<td>Turkey</td>
<td>2250</td>
<td>50.26</td>
<td>176</td>
<td>6.19</td>
<td>84.53</td>
<td>5633</td>
<td>70.96</td>
<td>31</td>
<td>2.22</td>
<td>51.23</td>
</tr>
<tr>
<td>Club 2</td>
<td>1843</td>
<td>52.24</td>
<td>108</td>
<td>6.63</td>
<td>91.04</td>
<td>8308</td>
<td>72.31</td>
<td>20</td>
<td>2.51</td>
<td>55.51</td>
</tr>
<tr>
<td>Argentina</td>
<td>7838</td>
<td>64.88</td>
<td>60</td>
<td>3.09</td>
<td>57.04</td>
<td>10170</td>
<td>74.45</td>
<td>15</td>
<td>2.34</td>
<td>57.77</td>
</tr>
<tr>
<td>Australia</td>
<td>10815</td>
<td>70.73</td>
<td>20</td>
<td>3.27</td>
<td>62.77</td>
<td>27872</td>
<td>80.44</td>
<td>5</td>
<td>1.76</td>
<td>48.43</td>
</tr>
<tr>
<td>Austria</td>
<td>8444</td>
<td>68.74</td>
<td>32</td>
<td>2.78</td>
<td>51.90</td>
<td>27567</td>
<td>79.03</td>
<td>5</td>
<td>1.38</td>
<td>46.97</td>
</tr>
<tr>
<td>Barbados</td>
<td>7039</td>
<td>65.87</td>
<td>61</td>
<td>4.26</td>
<td>81.20</td>
<td>15707</td>
<td>76.09</td>
<td>12</td>
<td>1.50</td>
<td>39.02</td>
</tr>
<tr>
<td>Belgium</td>
<td>8070</td>
<td>70.10</td>
<td>27</td>
<td>2.66</td>
<td>55.03</td>
<td>25264</td>
<td>78.36</td>
<td>4</td>
<td>1.64</td>
<td>52.26</td>
</tr>
<tr>
<td>Canada</td>
<td>10576</td>
<td>71.07</td>
<td>26</td>
<td>3.61</td>
<td>69.61</td>
<td>27845</td>
<td>79.92</td>
<td>5</td>
<td>1.52</td>
<td>44.43</td>
</tr>
<tr>
<td>Chile</td>
<td>5086</td>
<td>56.90</td>
<td>109</td>
<td>5.28</td>
<td>79.45</td>
<td>12141</td>
<td>77.91</td>
<td>8</td>
<td>1.99</td>
<td>49.18</td>
</tr>
<tr>
<td>Colombia</td>
<td>2819</td>
<td>56.68</td>
<td>92</td>
<td>6.76</td>
<td>98.11</td>
<td>6094</td>
<td>71.82</td>
<td>20</td>
<td>2.44</td>
<td>54.90</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>4513</td>
<td>61.45</td>
<td>81</td>
<td>7.22</td>
<td>98.05</td>
<td>8586</td>
<td>78.21</td>
<td>10</td>
<td>2.26</td>
<td>51.92</td>
</tr>
<tr>
<td>Denmark</td>
<td>11438</td>
<td>72.16</td>
<td>20</td>
<td>2.59</td>
<td>55.82</td>
<td>27970</td>
<td>77.42</td>
<td>5</td>
<td>1.76</td>
<td>51.28</td>
</tr>
<tr>
<td>Finland</td>
<td>7785</td>
<td>68.52</td>
<td>19</td>
<td>2.58</td>
<td>60.28</td>
<td>23784</td>
<td>78.48</td>
<td>4</td>
<td>1.76</td>
<td>49.90</td>
</tr>
<tr>
<td>France</td>
<td>8531</td>
<td>70.48</td>
<td>25</td>
<td>2.85</td>
<td>61.31</td>
<td>25664</td>
<td>79.72</td>
<td>4</td>
<td>1.88</td>
<td>53.14</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Country</th>
<th>GDP60</th>
<th>LIF60</th>
<th>IMR60</th>
<th>TFR60</th>
<th>DEP60</th>
<th>GDP03</th>
<th>LIF03</th>
<th>IMR03</th>
<th>TFR03</th>
<th>DEP05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greece</td>
<td>4177</td>
<td>68.73</td>
<td>50</td>
<td>2.20</td>
<td>53.24</td>
<td>15785</td>
<td>78.43</td>
<td>8</td>
<td>1.28</td>
<td>48.28</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>3322</td>
<td>66.32</td>
<td>33</td>
<td>5.31</td>
<td>77.60</td>
<td>27658</td>
<td>81.60</td>
<td>4</td>
<td>0.94</td>
<td>37.18</td>
</tr>
<tr>
<td>Iceland</td>
<td>8380</td>
<td>73.37</td>
<td>17</td>
<td>3.94</td>
<td>74.26</td>
<td>26347</td>
<td>81.06</td>
<td>3</td>
<td>2.00</td>
<td>51.12</td>
</tr>
<tr>
<td>Ireland</td>
<td>5294</td>
<td>69.69</td>
<td>28</td>
<td>3.98</td>
<td>73.23</td>
<td>28248</td>
<td>77.88</td>
<td>5</td>
<td>1.97</td>
<td>46.73</td>
</tr>
<tr>
<td>Israel</td>
<td>6750</td>
<td>68.70</td>
<td>29</td>
<td>3.85</td>
<td>69.28</td>
<td>20713</td>
<td>79.81</td>
<td>5</td>
<td>2.89</td>
<td>61.34</td>
</tr>
<tr>
<td>Italy</td>
<td>7167</td>
<td>69.31</td>
<td>40</td>
<td>2.50</td>
<td>51.69</td>
<td>22925</td>
<td>80.00</td>
<td>5</td>
<td>1.29</td>
<td>50.89</td>
</tr>
<tr>
<td>Jamaica</td>
<td>3477</td>
<td>64.28</td>
<td>61</td>
<td>5.64</td>
<td>85.11</td>
<td>4585</td>
<td>72.12</td>
<td>14</td>
<td>2.60</td>
<td>64.36</td>
</tr>
<tr>
<td>Japan</td>
<td>4509</td>
<td>67.91</td>
<td>25</td>
<td>2.02</td>
<td>56.12</td>
<td>24037</td>
<td>81.91</td>
<td>3</td>
<td>1.29</td>
<td>50.67</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>12920</td>
<td>68.81</td>
<td>29</td>
<td>2.37</td>
<td>47.42</td>
<td>49262</td>
<td>78.22</td>
<td>5</td>
<td>1.67</td>
<td>48.58</td>
</tr>
<tr>
<td>Mauritius</td>
<td>3662</td>
<td>59.63</td>
<td>61</td>
<td>5.72</td>
<td>96.35</td>
<td>16464</td>
<td>72.03</td>
<td>15</td>
<td>1.90</td>
<td>44.95</td>
</tr>
<tr>
<td>Mexico</td>
<td>3719</td>
<td>56.95</td>
<td>88</td>
<td>6.82</td>
<td>96.98</td>
<td>7938</td>
<td>75.07</td>
<td>20</td>
<td>2.37</td>
<td>57.70</td>
</tr>
<tr>
<td>Netherlands</td>
<td>10462</td>
<td>73.24</td>
<td>16</td>
<td>3.17</td>
<td>63.93</td>
<td>26157</td>
<td>78.81</td>
<td>5</td>
<td>1.72</td>
<td>48.28</td>
</tr>
<tr>
<td>New Zealand</td>
<td>12063</td>
<td>70.92</td>
<td>21</td>
<td>4.02</td>
<td>71.02</td>
<td>22195</td>
<td>79.30</td>
<td>6</td>
<td>1.97</td>
<td>50.64</td>
</tr>
<tr>
<td>Norway</td>
<td>9473</td>
<td>73.41</td>
<td>17</td>
<td>2.90</td>
<td>58.73</td>
<td>34011</td>
<td>79.37</td>
<td>4</td>
<td>1.81</td>
<td>52.12</td>
</tr>
<tr>
<td>Panama</td>
<td>2499</td>
<td>60.65</td>
<td>63</td>
<td>5.92</td>
<td>89.61</td>
<td>8244</td>
<td>74.81</td>
<td>20</td>
<td>2.69</td>
<td>57.11</td>
</tr>
<tr>
<td>Paraguay</td>
<td>2510</td>
<td>63.77</td>
<td>62</td>
<td>6.55</td>
<td>104.30</td>
<td>4716</td>
<td>70.85</td>
<td>35</td>
<td>3.44</td>
<td>68.34</td>
</tr>
<tr>
<td>Portugal</td>
<td>3689</td>
<td>63.34</td>
<td>76</td>
<td>3.07</td>
<td>59.08</td>
<td>17334</td>
<td>77.31</td>
<td>5</td>
<td>1.45</td>
<td>48.28</td>
</tr>
<tr>
<td>Singapore</td>
<td>4219</td>
<td>64.54</td>
<td>30</td>
<td>4.93</td>
<td>82.79</td>
<td>26999</td>
<td>78.91</td>
<td>3</td>
<td>1.34</td>
<td>38.96</td>
</tr>
<tr>
<td>Spain</td>
<td>4881</td>
<td>69.18</td>
<td>42</td>
<td>2.89</td>
<td>55.37</td>
<td>20644</td>
<td>80.09</td>
<td>4</td>
<td>1.30</td>
<td>45.45</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>866</td>
<td>61.97</td>
<td>63</td>
<td>5.50</td>
<td>85.97</td>
<td>4272</td>
<td>71.02</td>
<td>12</td>
<td>2.00</td>
<td>44.22</td>
</tr>
<tr>
<td>Sweden</td>
<td>11065</td>
<td>73.11</td>
<td>15</td>
<td>2.32</td>
<td>51.45</td>
<td>26136</td>
<td>80.18</td>
<td>3</td>
<td>1.67</td>
<td>52.99</td>
</tr>
<tr>
<td>Switzerland</td>
<td>15253</td>
<td>71.27</td>
<td>20</td>
<td>2.51</td>
<td>50.79</td>
<td>28792</td>
<td>80.82</td>
<td>4</td>
<td>1.41</td>
<td>47.34</td>
</tr>
<tr>
<td>Trinidad T</td>
<td>6274</td>
<td>63.52</td>
<td>48</td>
<td>4.99</td>
<td>88.72</td>
<td>18417</td>
<td>69.03</td>
<td>15</td>
<td>1.60</td>
<td>40.16</td>
</tr>
<tr>
<td>UK</td>
<td>10323</td>
<td>70.64</td>
<td>22</td>
<td>2.81</td>
<td>53.71</td>
<td>26046</td>
<td>78.57</td>
<td>5</td>
<td>1.71</td>
<td>51.62</td>
</tr>
<tr>
<td>USA</td>
<td>12892</td>
<td>69.88</td>
<td>25</td>
<td>3.31</td>
<td>66.72</td>
<td>34875</td>
<td>77.52</td>
<td>7</td>
<td>2.04</td>
<td>49.42</td>
</tr>
<tr>
<td>Uruguay</td>
<td>6143</td>
<td>67.71</td>
<td>48</td>
<td>2.90</td>
<td>56.25</td>
<td>8855</td>
<td>75.44</td>
<td>14</td>
<td>2.19</td>
<td>59.55</td>
</tr>
<tr>
<td>Club 3</td>
<td>7077</td>
<td>67.06</td>
<td>42</td>
<td>3.92</td>
<td>69.74</td>
<td>20798</td>
<td>77.42</td>
<td>9</td>
<td>1.86</td>
<td>50.41</td>
</tr>
<tr>
<td>Full Sample</td>
<td>3995</td>
<td>54.63</td>
<td>100</td>
<td>5.51</td>
<td>79.96</td>
<td>11246</td>
<td>68.21</td>
<td>39</td>
<td>3.24</td>
<td>64.93</td>
</tr>
</tbody>
</table>

142