Did Raising Doing Business Scores Boost GDP?

Tamanna Adhikari* Karl Whelan†

April 2023

Abstract

We use the time series variation in the World Bank's "distance to frontier" estimates of the ease of doing business to assess the effects of changes in this variable on real GDP per capita. The use of Vector Autoregression techniques allows us to identify shocks to the *Doing Business* scores that are initially uncorrelated with GDP, thus addressing an important endogeneity problem that affects the cross-sectional literature on this topic. We report a robust finding that improvements in *Doing Business* scores have at least a temporary negative impact on GDP and find little evidence for a positive effect in the years following these improvements.

*Central Bank of Ireland: tamannaadhikari2@gmail.com. All opinions expressed here are those of the authors and do not reflect the opinions of the Central Bank of Ireland

†University College Dublin. karl.whelan@ucd.ie. Corresponding author.

1. Introduction

In the huge literature on the determinants of economic development, the key goal for economists has been uncovering the policies and institutions that can help a country to achieve high levels of GDP per capita and living standards. In some cases, historical events such as the divisions of Germany and Korea have provided natural experiment that illustrate some of the key ingredients for successful economic development. Examples of this kind have generated widespread agreement among economists that economic development requires an environment in which private businesses can operate relatively freely. The exact roles played by private property rights, free markets, effective regulation and democracy are the subject of vigorous ongoing debates but it is clear that development agencies such as the World Bank and IMF place great weight on the idea that a business-friendly environment is important for promoting economic growth. Commentary about the need for "structural reforms" to boost growth features in many of their publications.

Assessing the link between a country's business environment and its level of economic development has always been difficult because of the absence of comparable cross-country data on the policy and regulatory environments that businesses operate in. However, the World Bank's *Doing Business* report, launched in 2003, offered the possibility of this gap being filled by providing a wealth of information on the business environment, summarised via a set of rankings. The annual publication stemmed from a series of academic studies sponsored by the World Bank examining the costs of starting a business and the ability to enforce contracts (see Djankov, La Porta, Lopes-de-Silanes, and Shleifer, 2002, and Djankov, La Porta, Lopez-de-Silanes, and Shleifer, 2003). This was followed over time by further detailed studies of other aspects of the business environment such as the ability to obtain credit, the efficiency of insolvency procedures, the costs and delays associated with exporting and importing and the efficiency of tax collection.[1] The data released by *Doing Business* were subsequently used in a wide range of research papers, most focusing on the sub-components of the overall index, such as measures of the ease of international trade or ease of paying taxes.[2]

Despite its widespread use by researchers, a number of questions emerged over time in relation to *Doing Business*. First, there was the question of whether the rankings were being "gamed" by governments. Second, there were concerns that the methodology was not always measuring the business environment experienced by typical firms. Finally, there were some high-profile controversies, most notably when it emerged in 2020 that there had been irregularities in the recording of data for some countries, with senior World Bank officials apparently yielding to political pressure

[1] See Djankov, McLiesh, and Shleifer (2007), Djankov, Hart, McLiesh, and Shleifer (2008), Djankov, Freund, and Pham (2010) and Djankov, Ganser, McLiesh, Ramalho, and Shleifer (2010).

[2] To give just a few examples, Djankov, Ganser, McLiesh, Ramalho, and Shleifer (2010) use the tax component of the survey to uncover negative effects of high corporate tax rates on entrepreneurship and investment. Lawless (2013) finds the complexity of the corporate tax system has a negative effect on inward foreign direct investment. Norback, Persson and Douhan (2014) find that better scores on the survey's measures of the ease of starting a business are associated with higher openness to trade while Freund and Bolasky (2008) find the benefits to economic development from openness are higher in economies with good scores on this measure.

to change scores. While the impact on these irregularities on the overall rankings were minor, the damage to the reputation of the project was severe. In September 2021, the World Bank announced that it was discontinuing the publication.[3]

This still leaves an important question. Did countries that improved their *Doing Business* scores tend to see benefits in terms of increases in GDP per capita? In this paper, we attempt to answer this question, making a number of new contributions.

Our first contribution is the use of a panel data set to explore the time series dimension of the relationship between *Doing Business* scores and real GDP per capita. Because of the relatively short time series available from the *Doing Business* survey, existing research on this question has focused on pure cross-sectional relationships. For example, Djankov, McLiesh and Ramalho (2006) find that countries with better *Doing Business* scores have consistently grown faster while Gillanders and Whelan (2014) find that the ease of doing business rank works better as an explanatory variable for GDP per capita than other popular institutional variables such as the rule of law index from the World Bank's Governance Matters dataset, which has been a popular explanatory variable in the literature on the link between institutions and development. These papers, however, relied purely on cross-sectional data to identify these relationships. With fifteen years of data from the survey now available, we believe there is sufficient time dimension in the data now to warrant an investigation of the impact of variation over time in the *Doing Business* indicators.

The use of panel data has a number of advantages over cross-sectional analysis. It can potentially expose relationships that exist in the cross-section as being spurious. In addition, it is unlikely that cross-sectional relationships can provide useful estimates of the short- or medium-term impact of policy changes. For example, if a poor country with significant restrictions on business introduces a major program of structural reforms, it is highly unlikely that GDP would immediately jump up to the level predicted by purely cross-sectional estimates. A time series approach is required to assess the dynamics of this relationship.

Our second contribution is to use a new approach to identifying the effect of *Doing Business* scores on GDP. It is well understood that the cross-sectional literature on the impact of policies and institutions on economic development is plagued by endogeneity problems. We may be able to identify a set of good policies that possibly encourage high levels of economic development but high-income countries are also more likely to have a social and political consensus that favors such policies. Much of the previous research on this topic has attempted to deal with this problem using instrumental variables such as geographic or climate-related variables (as in Hall and Jones, 1999) or information on a country's colonial past (as in Acemoglu, Johnson and Robinson, 2001) to identify plausibly exogenous variations in variables that have had a likely influence on institutions and economic policy.

[3] https://www.worldbank.org/en/news/statement/2021/09/16/world-bank-group-to-discontinue-doing-business-report

This is the approach followed by Djankov, McLiesh and Ramalho (2006) and Gillanders and Whelan (2014) when analyzing the impact of the ease of doing business on GDP per capita. While this type of research has been influential, it is affected by problems due to the limited amount of variation in institutions associated with the instruments (and thus weak first-stage fits), controversies over the true exogeneity of various instruments and collinearity between different plausible institutional or policy variables. Kelly (2021) also suggests many apparently significant findings in the cross-sectional literature on economic development may reflect spurious spatial noise.

In contrast to the cross-sectional approach, we use Vector Autoregression (VAR) techniques to identify the dynamic impact of shocks to the ease of doing business that are, by construction, uncorrelated with GDP during the period when they occur, thus avoiding the endogeneity issue that has been the principal focus of the cross-sectional literature. Because we would expect regulatory reforms to take a number of years to have their full impact, we would still expect the shocks we identify to have an impact on GDP over time and these impacts can be assessed using a VAR.

A final contribution is the use of a new time series to reflect *Doing Business* scores. Most previous research in this area has used the *Doing Business* ranking as its explanatory variable. However, the ranking is constructed from a set of underlying scores and movements in a country's rank may reflect developments elsewhere rather than within the country. We also document that relatively small changes in the underlying indicators can result in large changes in a country's rank, depending on where the country sits in the overall distribution. In recent years, however, the *Doing Business* team published a series of "distance to frontier" scores for the underlying indicators used to construct the ease of doing business rank. We construct new time series of *Doing Business* scores using this data to provide a series constructed according to a single methodology that is available for the years 2006-2020. We believe this variable is more suitable than the rank variable for both cross-sectional analysis as well as the dynamic analysis that is undertaken in this paper.

Our results are perhaps somewhat surprising. We do not find evidence of a consistent positive effect of improvements in the *Doing Business* score. In fact, we report that higher *Doing Business* scores appear to have at least a temporary negative impact on GDP. This finding is robust across many different specifications and holds for different sources of GDP data. We obtain similar (though somewhat weaker) findings when we look at the sub-components of the overall *Doing Business* score.

The paper is organized as follows. Section 2 discusses the *Doing Business* project, outlines some of the criticisms and controversies associated with it and presents summary statistics on the data that we use. Section 3 discusses theoretical considerations and presents our results. Section 4 provides some conclusions and discusses possible explanations for the results.

2. Data

In this section, we briefly describe the *Doing Business* project in more detail and present the data from the project that we use.

2.1. The Doing Business Project: Origins and Controversies

The origins of the *Doing Business* project have been described by Djankov (2016). In the early 2000s, the World Bank had been running a project on the importance of institutions in development, at the instigation of then chief economist Joseph Stiglitz. World Bank economist Simeon Djankov turned to a research team lead by Andrei Shleifer to co-operate with the World Bank to produce a series of papers measuring business-related institutions and policies around the world. The work from these early papers, such as Djankov, La Porta, Lopes-de-Silanes, and Shleifer (2002, 2003) constructed a series of data sets that resulted in the first *Doing Business* rankings based on five indicators: starting a business, enforcing contracts, getting credit, employing workers and resolving insolvency. In subsequent years, a number of other studies were done and the most rankings were also based on six additional indicators: dealing with construction permits, registering property, protecting minority investors, paying taxes, trading across borders and getting electricity.

While the project could be considered a success from a research perspective, with the underlying papers all published in leading journals and the data set used extensively by other researchers, it also created some controversy. The "Employing Workers" component of the survey was heavily criticized by international labor organizations on the grounds that the adoption of potentially exploitative labor practices could boost a country's ease of doing business rank. The World Bank responded to this criticism in 2009 by removing the employing workers measure from the overall rankings but concerns remained. For example, in 2018, the International Trade Union Confederation argued the program should be scrapped, claiming "*It is entirely inappropriate for the World Bank, as a publicly-funded multilateral institution, to endorse and promote a right-wing platform of weakened regulations and reduced taxation on business around the world.*"[4]

Another concern related to the methodology underlying the survey. The project relied on contacting lawyers, accountants, architects, engineers and other professionals to get their estimates of the time and costs for a typical business in the largest business city in a country associated with undertaking various tasks.[5] World Bank (2017) reported that about 60 percent of the data collected for the project were based on a reading of the law. In this sense, the team behind the project argued that the data collected were largely objective and factual. The *Doing Business* team argued this ap-

[4] https://www.ituc-csi.org/world-bank-should-scrap

[5] Besley (2016) describes the typical firm assumptions in more detail. "The central case is a firm with at least 60 employees, which is located in the country's largest business city. It is a private, limited-liability company and does not operate in an export-processing zone or an industrial estate with special export or import privileges. It is 100 percent domestically owned, and exports constitute more than 10 percent of its sales."

proach was superior to surveying firms directly because they got more objective information from contacting specialist professionals, particularly as many of the issues being surveyed, such as getting hooked up to the electricity grid or going though an insolvency process, were things that only happened very occasionally for any individual firm. Still, Hallward-Driemeier and Pritchett (2015) used evidence from the World Bank's Enterprise Surveys of individual firms to show that there was a limited correlation between some of the measures produced by the *Doing Business* project and the experiences actually reported by firms while Kar, Pritchett, Roy and Sen (2019) suggest this correlation gets weaker for states with weak governance institutions. This points to potentially serious weaknesses with the process of data collection for *Doing Business*.

A final concern about *Doing Business* was that it may have incentivized governments to "game" the indicators, focusing their reform efforts on items that are measured by the World Bank but which have a limited substantive impact on the economy. For examples, the 2020 report shows Rwanda with a rank of 38th in the world for ease of doing business, ahead of the Netherlands in 42nd. This may reflect gaming of the indicators rather than it being easier to do business in Rwanda than the Netherlands. While gaming of this sort may be harmless, some have argued that the project has been damaging because it can focus government energies on the wrong kinds of reforms. For example, Arrunada (2007) reviews the regulations and problems associated with starting a business and argues that *Doing Business* is "*promoting sterile reforms while speeding up useless formalities*."

The concerns and criticisms noted here may be worth keeping in mind when considering the potentially puzzling empirical results that we obtain in this paper

2.2. Data on Ease of Doing Business

Underlying the well-known rankings for the ease of doing business are a set of scores relating to 50 different indicators from the ten categories listed above. For each of these indicators, the scores run from 0 to 100, with 100 being the best possible value. The scores are constructed on a "distance to frontier" approach, so a score of 100 represents the best value for an indicator in the dataset. The underlying indicators are averaged for each of the ten categories and then an average is taken across the ten categories. It is this aggregate score that is used to construct the rankings. The World Bank *Doing Business 2017* report gives a description of how distance to frontier is constructed for one of the categories, Getting Electricity:

> *For the getting electricity indicators, for example, the frontier is set at three procedures, 18 days and no cost to obtain a new electricity connection in the economy's largest business city. The worst for the same group of indicators is set at 9 procedures, 248 days and 81 times the economy's income per capita as the cost. In addition, the getting electricity indicators measure the reliability of electricity supply and transparency of tariffs through an index ranging from 0 to 8; in this case 8 is the frontier score. For example, in the case of reliability and transparency, an economy with a*

score of 6 would be considered to be 75% of the way to the frontier and would have a distance to frontier score of that value.

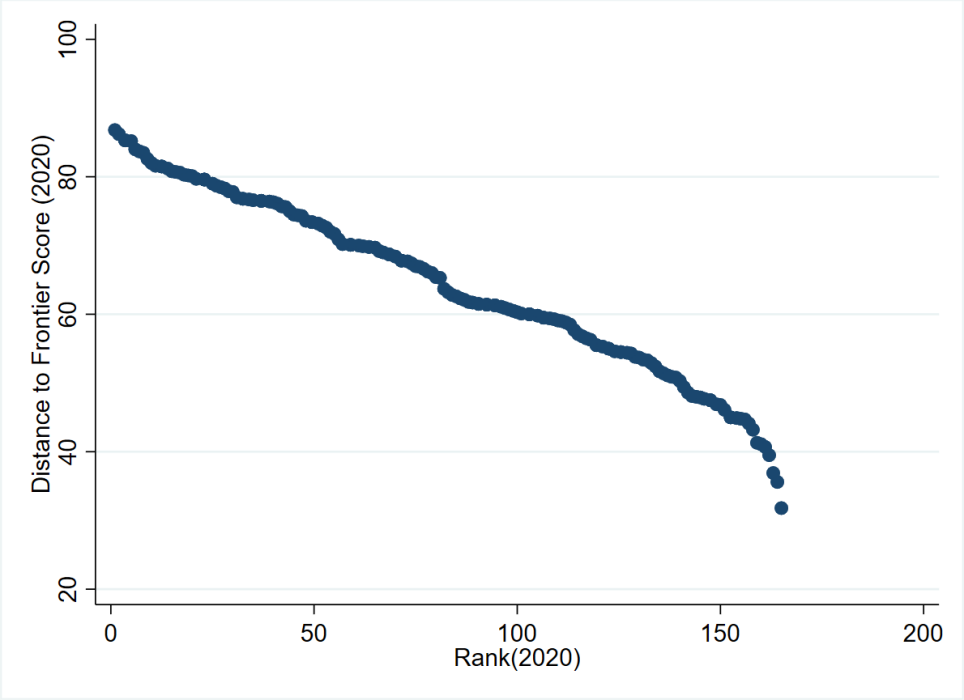
In recent years, the *Doing Business* team made available all of the underlying distance to frontier (DTF) data as well as the aggregated DTF scores for each country. To give a sense of the magnitudes of the aggregated DTF scores, the average aggregated DTF score for 2020 was 64.14 and the standard deviation was 12.46.

Figure 1 illustrates the relationship between the underlying DTF score and the rank using 2020 data. Note that despite its name, the scoring system adopted by the World Bank means that a high "distance to frontier" score represents a good performance rather than a large distance from the frontier. The figure shows the relationship is relatively non-linear in places so a gain or drop of a particular size in the DTF score can translate into quite different changes in the rank depending on where a country starts out. For example, New Zealand was ranked first for ease of doing business in the 2020 report with a DTF score of 87.01. If New Zealand's DTF score had fallen by 5, its rank would have dropped by 9 places. In contrast, Sri Lanka was ranked 99th with a DTF score of 61.8. If Sri Lanka's DTF score had fallen by 5, its rank would have dropped 26 places. We believe this is an argument in favor of using the underlying DTF scores in empirical analysis rather than the rank variable.

Another factor in favor of using the DTF score is that changes in a country's DTF score will be almost completely due to changes in that particular country and not changes elsewhere. Changes to the frontier scores do not play a major role in year-to-year changes in these scores: According to the World Bank both the best and the worst performance were established every five years based on the Doing Business data for the year for which they were established and remained at that level for the next five years. In contrast, a country's ease of doing business rank could move around from year to year due to developments elsewhere rather than factors related to its DTF score or underlying changes in the business environment. As we will see, also, the average DTF scores reported by the survey have improved over time. This means that countries where the underlying scores were not changing were likely to see their rankings worsen over time.

There are, however, two practical problems when seeking to use the data on country-level aggregated DTF scores. The first is that the World Bank only published the aggregate DTF score from 2010 onwards, even though the DTF scores for the underlying indicators are available going back to as far as 2004 for some. The other problem is that the methodology used by the World Bank to put together the aggregate scores has changed a number of times because new sub-indicators have been added over time. So, for example, in our data set, there are only six years available for scores constructed according to the most recent methodology.

Figure 1: DTF Score vs Doing Business Rank



To obtain an aggregated DTF score for each country that allows us to use a full panel with the same number of observations as the underlying scores, we constructed a DTF variable that mimics the methodology the World Bank used to construct the aggregate DTF from 2015 to 2020 inclusive. Specifically, we regressed the aggregate DTF scores as published by the World Bank on nine of the category variables that are available widely during this period (we excluded the Getting Electricity category because of limited data availability). The results of this regression are shown in Table 1. The fitted values from the regression produces a measure that essentially replicates the current global DTF scores, with the R^2 for the regression being about 0.98. We then used the coefficients from this regression to construct a DTF measure from the underlying category scores that is available from 2006 to 2020. This is the explanatory variable that we use in our analysis.

Summary statistics for this measure are shown in Table 2 while Figure 2 shows a histogram of all of values in our sample. An alternative approach of equally weighting each of the underlying categories produces essentially identical results to those we report here.

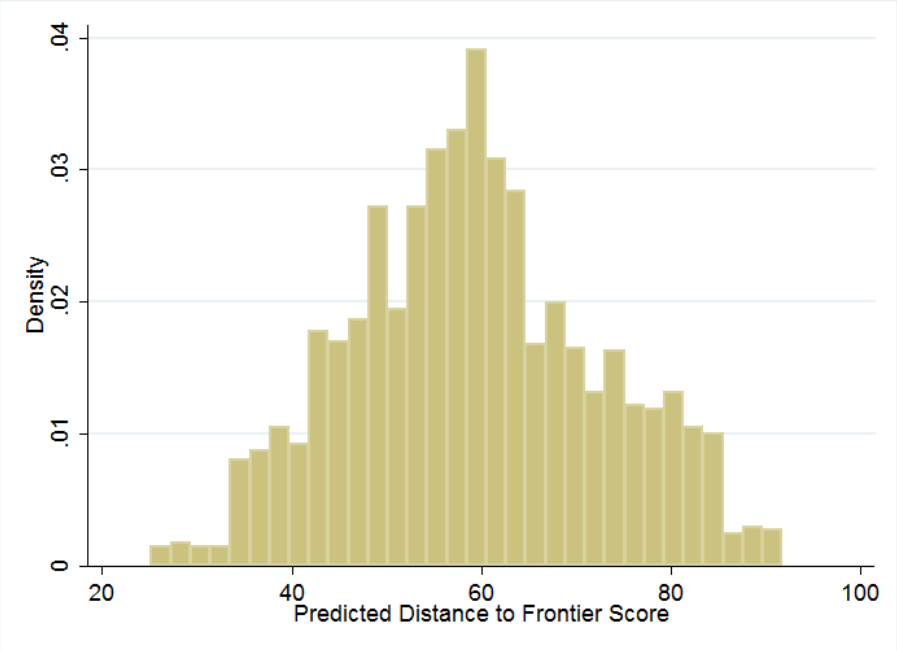
Table 1: Dependent Variable: Distance to Frontier

Variable	**Coefficient**
	(Std. Err.)
Starting a Business	0.099***
	(0.002)
Dealing with Construction Permits	0.130***
	(0.002)
Registering Property	0.117***
	(0.003)
Getting Credit	0.110***
	(0.001)
Protecting Minority Investors	0.105***
	(0.003)
Paying Taxes	0.116***
	(0.002)
Trade across Borders	0.108***
	(0.002)
Enforcing Contracts	0.104***
	(0.004)
Resolving Insolvency	0.092***
	(0.002)
Intercept	1.333***
	(0.341)
N	1815
R^2	0.98

Table 2: Predicted Distance to Frontier by Year

Year	Mean	N
2006	54.02	165
2007	55.13	165
2008	56.15	165
2009	57.29	165
2010	58.56	165
2011	59.18	165
2012	60.25	165
2013	60.91	165
2014	62.57	165
2015	63.20	165
2016	61.94	165
2017	62.07	165
2018	62.537	165
2019	63.33	165
2020	64.14	165
Total	60.02	2475

Figure 2: Distribution of Distance to Frontier



2.3. Data on Real GDP

Our analysis assesses the relationship between the DTF variable and real GDP per capita. The principal measure of real GDP that we use is taken from the World Bank's World Development Indicators (WDI) database. These data are collected from national statistical organizations and central banks by visiting and resident World Bank missions while data for selected high-income economies are taken from the OECD.

We also report some results using real GDP from the Penn World Tables (PWT) as well as the World Bank's PPP-based measures. Our panel data analysis employs country-specific fixed effects, so our identification of the relationship is based on within-country variation. For this reason, the kind of cross-country real income comparisons for which the PWT measures (and the World Bank's PPP-based measures) were designed are less relevant to our analysis. Johnson et al (2013) also argue that the current PWT data are not well-suited to high-frequency analysis using annual data, which is the type of analysis that we are carrying out.

Figure 3 illustrates the relationship between distance to frontier and the WDI measure of real GDP per capita using data from 2020 while Figure 4 substitutes the ease of doing business rank for the DTF score. Figure 5 shows the relationship between the growth in GDP and DTF between 2006 and 2020. All scatter plots show the expected sign. The DTF score is positively correlated with real

GDP per capita and the *Doing Business* rank is negatively correlated. The relationship between the cumulative growth in DTF and growth in GDP is weaker but still statistically significant.

These relationships may reflect the positive impact that improvements in *Doing Business* scores have on GDP per capita but it is also possible that these relationships reflect reverse causality. This may happen because stronger economies with greater levels of prosperity can create interest groups that focus on protecting the business sector that is key to prosperity. Another possible reverse causality link is that the various professionals consulted by the World Bank to create the indexes tend to have a more positive views about more developed economies. Ultimately, to assess the causal link, we will need a method for identifying some exogenous variations in the *Doing Business* scores.

Figure 3: GDP Level vs Distance to Frontier Score

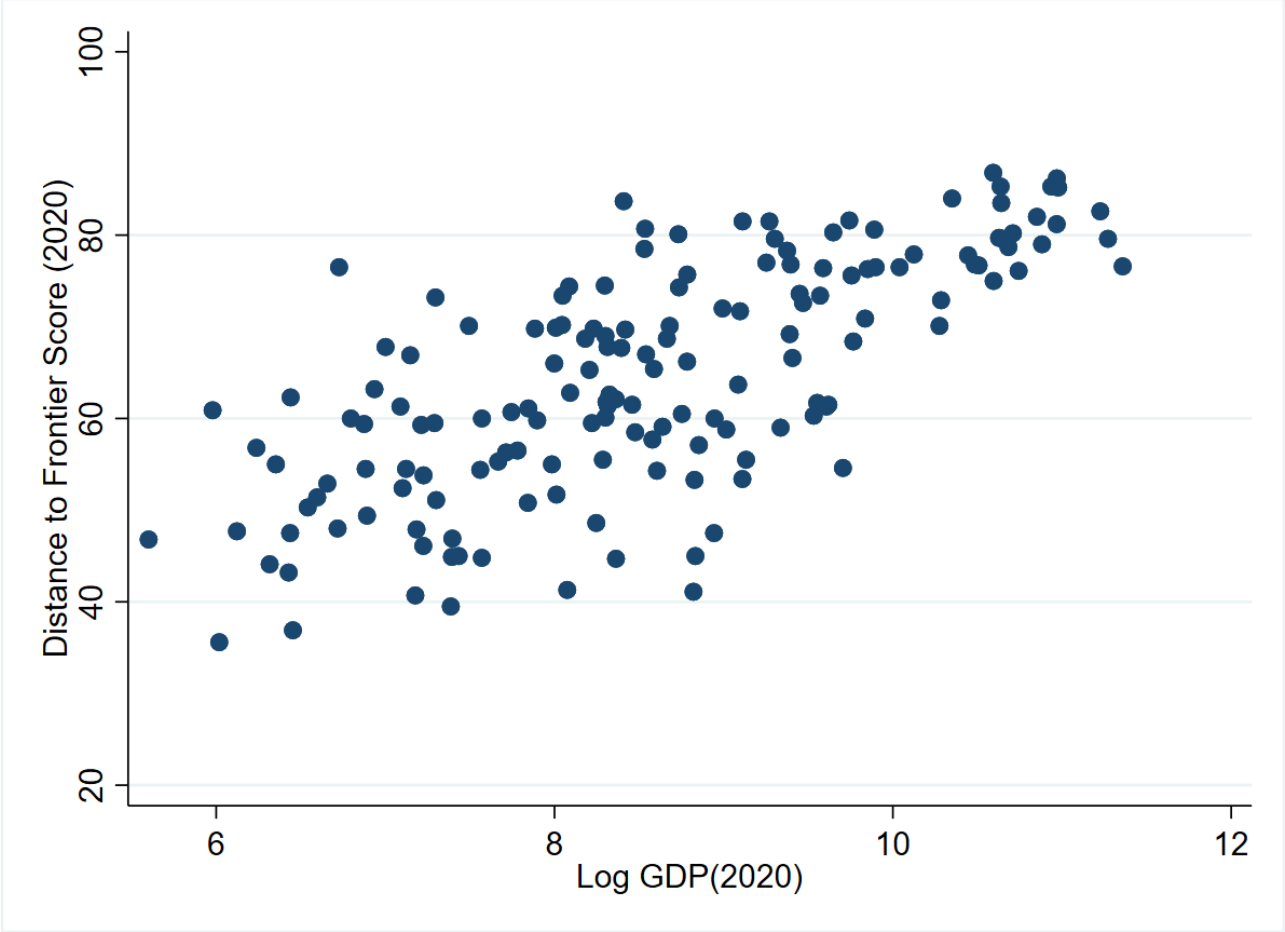


Figure 4: GDP Level vs Doing Business Rank

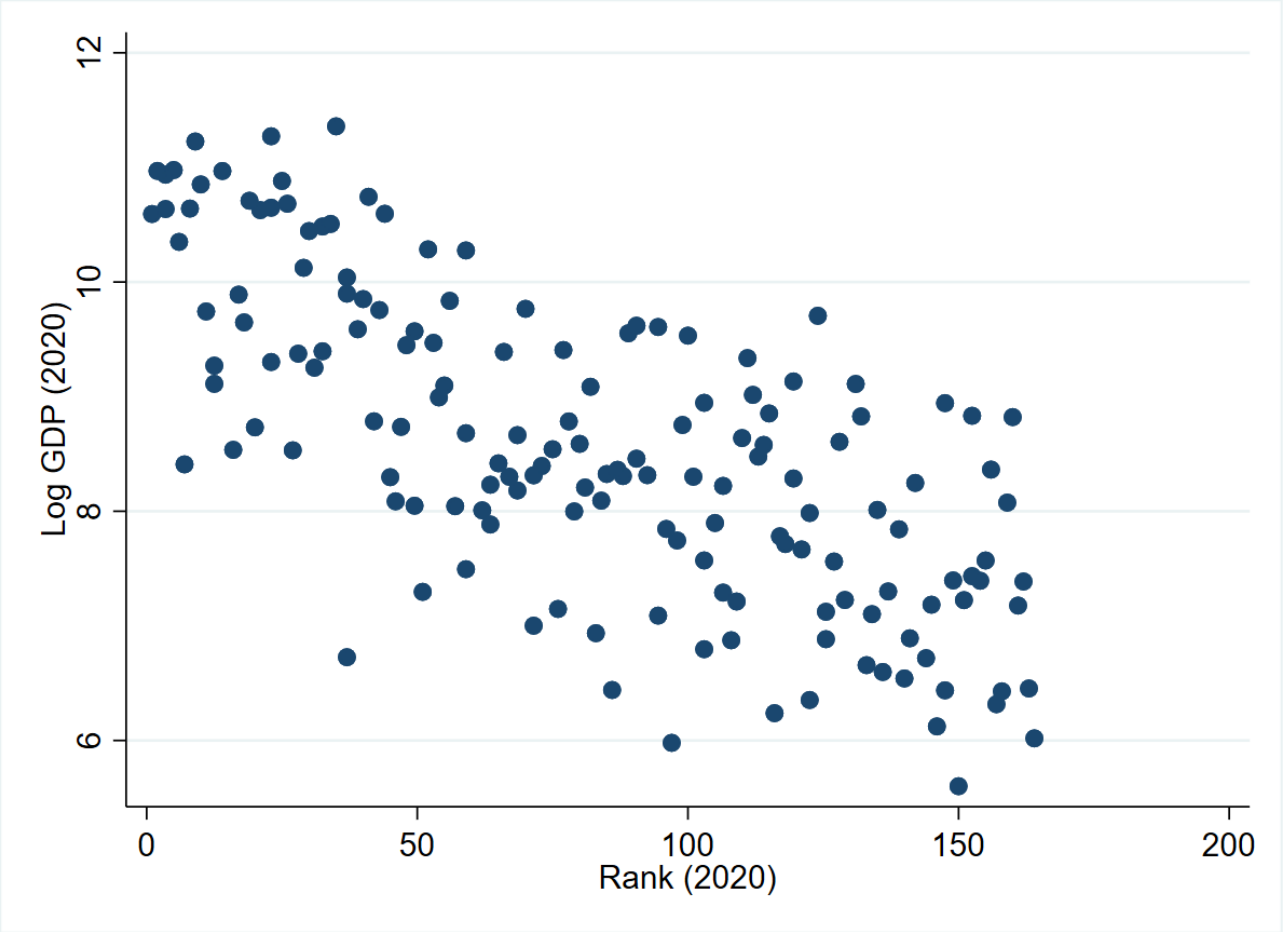
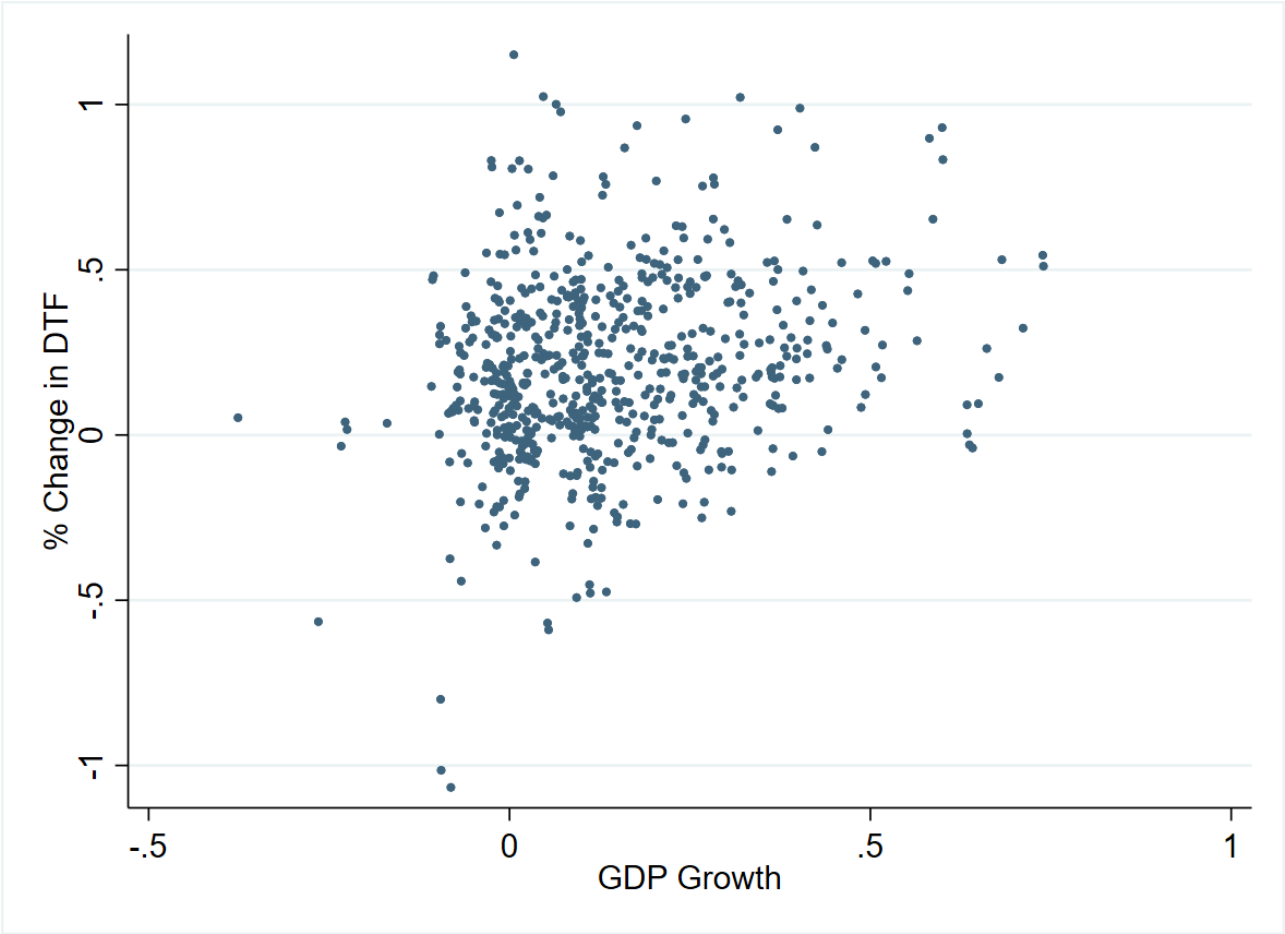


Figure 5: GDP Growth vs Distance to Frontier Score Growth



3. Theoretical Considerations and Empirical Evidence

Here, we briefly describe some theoretical consideration for what the relationship may be between real GDP per capita and our ease of doing business measure and then present our regression-based evidence.

3.1. Theoretical Considerations

From the perspective of the classic Solow (1956) model of economic growth, there are two channels through which restrictions on doing business could impact long run living standards. First, they can raise the efficiency which which the economy uses its resources, thus boosting total factor productivity. Second, they can raise investment in physical and human capital. The latter effect is well-known to boost the long-run level of GDP per worker rather than its growth rate. Similarly, the introduction of a single set of policies seems likely to induce a once-off improvement in total factor productivity rather than permanently boost its growth rate.

Of course, there are some models, such as Romer (1990), where policy changes can induce permanent changes in economic growth rates by inducing more investment in innovation. However, these models are probably best viewed as models of the determinants of global income levels or the GDP of the leading countries and are unlikely to be relevant for most of the countries in our sample. Of course, investment in innovation and in adapting technologies is important for countries that are not at frontier levels of total factor productivity but the appropriate model for most countries is probably more like the technology adoption model in Chapter 21 of Acemoglu (2008) based on Nelson and Phelps (1966). In that model, the level of a country's GDP per technological efficiency as a fraction of the leading country is determined by its ability to introduce leading technologies and an improvement in technology adaption can produce a "growth miracle" raising the growth rate for a while as the economy transitions to its new higher level of technology. But ultimately this is a level effect.

For these reasons, while we use a specification with the growth rate as a check, our main regression specification follows the standard approach in the panel data literature since Islam (1995) and models the impact of the DTF score on the level of GDP per capita with time and country effects proxying for the impact of technology.

3.2. Dynamic Panel Specification

The scatter plots just shown fit with intuition for why we might expect to find a relationship between real GDP per capita and the ease of doing business as measured by the DTF score. But a simple cross-sectional scatter plot is not a substitute for a careful econometric analysis. The cross-sectional relationship may be purely driven by endogeneity, with the causal direction going from GDP to the DTF measure. Alternatively, a relationship that looks strong in a cross-sectional analysis may break

down when we look at the time series variation within countries. In this section, we report our econometric analysis into this relationship. We use a balanced panel with 165 countries and with 15 years of available data (2006 to 2020) for each country.

We start by reporting results for the standard dynamic panel specification that has been used extensively in empirical growth analysis since Islam (1995). The specification allows for differences in the long-run levels of GDP for each county by including country-level fixed effects and controls for global economic developments by including time effects. We report a number of specifications including various lags of the log of real GDP and the log of the DTF variable. This gives us a specification of the form:

$$\begin{aligned} \log GDP_{i,t} &= \gamma_i + \mu_t + \beta_1 \log GDP_{i,t-1} + ... + \beta_n \log GDP_{i,t-n} \\ &\quad + \delta_0 \log DTF_{i,t} + ... + \delta_n \log DTF_{i,t-n} + \epsilon_{i,t} \end{aligned} \tag{1}$$

Because both variables have been logged, the coefficients can be interpreted as elasticities.

The results from estimation of this specification via OLS are shown in Table 3. Column 1 shows there is still a positive estimated relationship between DTF and real GDP per capita even after controlling for time and country-level fixed effects. This suggests that the relationship between the two variables is not simply a cross-sectional relationship and that the within-country variation in DTF, which is what is being used here to identify this coefficient, has the expected positive relationship with GDP. However, once we include some dynamics the picture is less clear. Including one lagged value of DTF of GDP per capita, we estimate a negative coefficient on the first lag of DTF which is larger than the contemporaneous positive coefficient. When second lagged values are added, we find a positive coefficient on the second lag of DTF. Third lags are not found to be statistically significant.

Table 4 summarizes the predictions for the long-run effect of DTF on GDP by reporting the same specifications re-written to show the sum of the coefficients on the DTF terms. Depending on the dynamic specification chosen, DTF has either a significant positive effect (no lags), a significant negative effect (one lag) or an insignificant negative effect (two lags).

Table 3: Estimating the Effect of DTF on Real GDP Per Capita

	(1)	(2)	(3)
	Per Capita GDP	Per Capita GDP	Per Capita GDP
GDP_{t-1}		0.945***	0.985***
		(0.012)	(0.023)
GDP_{t-2}			-0.0.061***
			(0.026)
DTF_t	0.319***	0.036***	0.054
	(0.0043)	(0.035)	(0.0045)
DTF_{t-1}		-0.059*	-0.180***
		(0.0042)	(0.064)
DTF_{t-2}			0.121***
			(0.044)
N	2475	2301	2136
F	1861	1217	1900

standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Estimating the Effect of DTF on Real GDP Per Capita

	(1)	(2)	(3)
	$\log \text{GDP}_t$	$\log \text{GDP}_t$	$\log \text{GDP}_t$
$\log GDP_{t-1}$		0.945***	0.985***
		(0.012)	(0.023)
$\log GDP_{t-2}$			-0.061***
			(0.026)
$\log DTF_t$	0.319 ***	-0.022***	-0.004
	(0.033)	(0.019)	(0.02)
$\Delta \log DTF_t$		0.059*	0.059
		(0.042)	(0.028)
$\Delta \log DTF_{t-1}$			-0.121
			(0.044)
N	2475	2301	2136
F	3754	18784	17996

Notes: Model estimated via OLS with time and country dummies.

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.3. VAR Specification

As with the positive cross-sectional relationship between real GDP per capita and the DTF score, one reason to potentially distrust the positive contemporaneous coefficients estimated in Table 3 is that there is likely to be an endogeneity issue. The higher the average living standard in a country is, the more likely it is to have a more powerful set of political and social forces favoring business-friendly policies and so any estimated positive relationship may simply reflect causality running from real GDP to the policies summarized by the DTF variable.

This endogeneity has been recognized widely in the literature that attempts to link economic policies and institutions with levels of GDP or long-run growth rates. However, this literature has focused largely on cross-sectional analysis, so the proposed solution to the endogeneity problem has been to construct exogenous instruments to assess the effects of the explanatory policy variable. In the case of the ease of doing business rank, for example, Djankov, McLiesh and Ramalho (2006) use instruments such as dummy variables for the origin of the country's legal institutions and principal religion as well dummy variables such as distance from the equator in their analysis of the impact of the doing business rank on long-run growth. Gillanders and Whelan (2014) use the following variables as instruments for the ease of doing business rank in their analysis of the relationship between the real and the level of real GDP per capita: Distance from the equator, a landlocked dummy, a measure of the suitability of the soil for rain fed crops and the proportion of land that is suitable for irrigated rice crops.

Within the confines of a purely cross-sectional analysis, it is clear that some kind of instrumental variable approach has to be adopted but the existing literature in this area has some important weaknesses. In many cases, the variables chosen as instruments have only very weak correlations with the policy variables being instrumented and these low-first stage fits are well-known to lead to unreliable inference.[6] The cross-sectional analysis is also, by definition, unable to answer important questions about the impacts over time of policy changes: If, for example, a country introduces a set of reforms that improve its ease of doing business scores, over what time period can it expect these changes to impact its GDP? The cross-section estimates are best viewed as giving the long-run effects with no information of the short- or medium-term dynamics.

To address these weakness, we employ a panel Vector Autoregression approach and report the impulse responses of real GDP per capita to shocks to the DTF score. The basic VAR specification consists of estimating two dynamic equations, one for the log of real GDP per capita and one for the

[6] See, for instance, Staiger and Stock (1997) and Lee et al (2021).

log of the DTF variable.

$$\begin{aligned}\log GDP_{i,t} &= \gamma_i^{GDP} + \mu_t^{GDP} + \beta_{11}\log GDP_{i,t-1} + ... + \beta_{1n}\log GDP_{i,t-n} \\ &\quad +\delta_{11}\log DTF_{i,t} + ... + \delta_{1n}\log DTF_{i,t-n} + \epsilon_{i,t}^{GDP} \quad (2)\\ \log DTF_{i,t} &= \gamma_i^{DTF} + \mu_t^{DTF} + \beta_{21}\log GDP_{i,t-1} + ... + \beta_{22n}\log GDP_{i,t-n} \\ &\quad +\delta_{21}\log DTF_{i,t} + ... + \delta_{2n}\log DTF_{i,t-n} + \epsilon_{i,t}^{DTF} \quad (3)\end{aligned}$$

Testing for lag lengths, we select the best-fitting specification as the one with 2 lags, meaning our VAR has an effective T of 13 but the key results reported here are also obtained with different lag lengths. Table 5 reports the coefficients from the two regressions that make up our estimated VAR.

The approach to identification of shocks in a VAR analysis relies on the assumptions made about the relationships between the reduced-form residuals $\epsilon_{i,t}^{GDP}$ and $\epsilon_{i,t}^{DTF}$ and an assumed underlying set of structural shocks. In our analysis, we follow the Cholesky decomposition approach and place the log of real GDP per capita first in the ordering of shocks. This means we are assuming the contemporaneous correlation between $\epsilon_{i,t}^{GDP}$ and $\epsilon_{i,t}^{DTF}$ is fully driven by a causal link going from GDP to DTF. With this assumption, we can then analyse the impact on GDP of that component of $\epsilon_{i,t}^{DTF}$ which is uncorrelated with $\epsilon_{i,t}^{GDP}$.

The assumption that contemporaneous correlation between $\epsilon_{i,t}^{GDP}$ and $\epsilon_{i,t}^{DTF}$ is fully driven by the causal link from GDP to DTF may seem strong but there is merit in focusing on the dynamic effects of those shocks that are uncorrelated with contemporaneous changes in GDP. The effects of changes in business regulations are likely to emerge over time rather than being quickly identifiable short run impacts. Moreover, a regression of $\epsilon_{i,t}^{DTF}$ on $\epsilon_{i,t}^{GDP}$ reveals a statistically insignificant relationship (a t-statistic of 1.25), so the structural DTF shocks we are examining still account for the vast majority of the random variation in the DTF series.

The estimated impulse-response functions for one standard deviation shocks of both types are shown as the black lines in Figure 6. The lines above and below are 10% and 90% confidence intervals generated using 1000 bootstrap replications of the estimated VAR model. We use bootstrapped standard errors because the time element of our sample is relatively small and error terms in this kind of cross-sectional dataset tend to be heteroskedastistic and non-normal, thus making alternatives such as asymptotic standard errors or Monte Carlo error bands less accurate.[7] Given the limited time series dimension of our dataset, we limit ourselves to showing estimated responses up to seven years after the initial shock as it would be unwise to make judgments on very long-run dynamics based on a relatively short dataset.

The left-hand graphs in Figure 6 show the response of GDP and DTF to one-standard deviation

[7]Each draw of the bootstrap residuals draws from the same historical time period so the bootstrapped residuals contain the same correlations across countries and for the two variables within countries as in the historical data.

shocks to GDP while the right-hand graphs show the responses to one-standard deviation shocks to DTF. As expected, a positive shock to GDP provides a boost to DTF over the time horizon we are examining. More interesting is the upper-right graph, which shows the dynamic impact on GDP of a shock to DTF. Since this is our main point of interest, we have shown this graph on its own as Figure 7. The results are fairly striking. By definition, the DTF shock has no impact on GDP in period zero (when the shock occurs) but it then has a negative impact that is estimated to only gradually fade away over time. In terms of magnitudes, we have multiplied the series by 100, so we can see from the bottom right chart in Figure 6 that the one-standard deviation shock to DTF is about 3 percentage points in size. The point estimate of the maximum reduction in GDP is estimated at 0.39 percentage points three periods after the shock, albeit the bootstrapped confidence intervals suggest a wide amount of uncertainty around this figure.

By definition, all shocks in this model are temporary, so the improvement in DTF modeled here fades away over time. However, one can use the estimated VAR coefficients to calculate the theoretical impact of a permanent shock to DTF of the same size as the initial shock modeled here. Figure 8 shows that a permanent improvement of 3 percentage points in DTF is estimated to be associated with a reduction in GDP of about 0.46 percentage points.

One concern about this finding is that perhaps the absence of any long run significant effect for DTF is the result a low-power research design due to short time period available. However, we have performed various checks and we do not believe this is the explanation. First, we have worked with versions of our econometric VAR model adapted so the DTF coefficients imply modest positive long-run on GDP per capita. Bootstrapped simulations of this model produce clearly statistically significant short and long-run impacts. In other words, in a model with our sample size and with this kind of sampling variability, our techniques would pick up a significant effect if it was there. Second, the sample is long enough to pick up other well known significant effects in the growth literature. For example, since at least De Long and Summers (1991), the investment share of GDP is robustly correlated with increases in GDP per capita. In an appendix, we report the results from replacing DTF in the VAR with the investment share of GDP. The responses are positive and statistically significant at all horizons, showing the small sample size does not prevent significant impacts from being detected.

Table 5: VAR Regressions

	(1)	(2)
	log GDP$_t$	log DTF$_t$
GDP_{t-1}	0.986 ***	0.018*
	(0.0023)	(0.011)
GDP_{t-2}	-0.060***	0.009
	(0.0268)	(0.013)
log DTF_{t-1}	-0.126***	0.992 ***
	(0.045)	(0.022)
log DTF_{t-2}	0.0114**	-0.124***
	(0.044)	(0.021)
N	2136	2136

Notes: Model estimated with time and country dummies.

Notes: Standard errors in parenthesis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 6: Impulse Responses to One Standard Deviation Shocks (Includes Bootstrapped 10% and 90% Confidence Intervals)

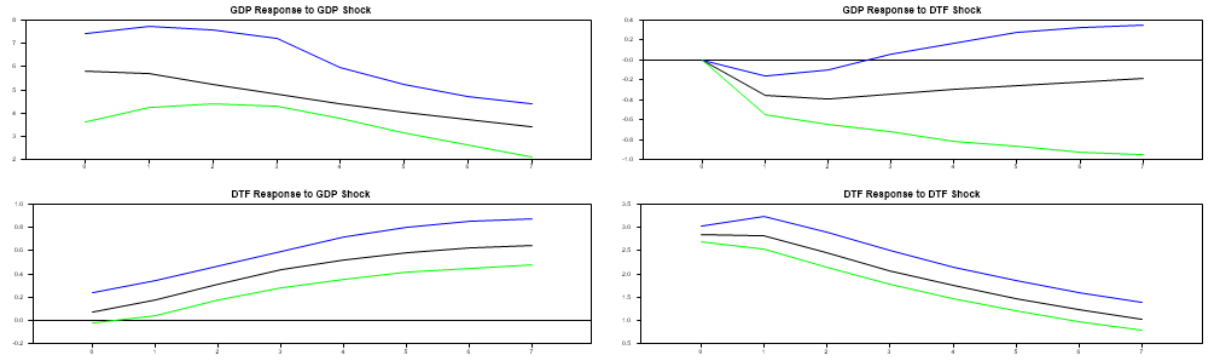


Figure 7: Response of the Level of GDP to a One Standard Deviation DTF Shock (Baseline Specification, Includes Bootstrapped 10% and 90% Confidence Intervals)

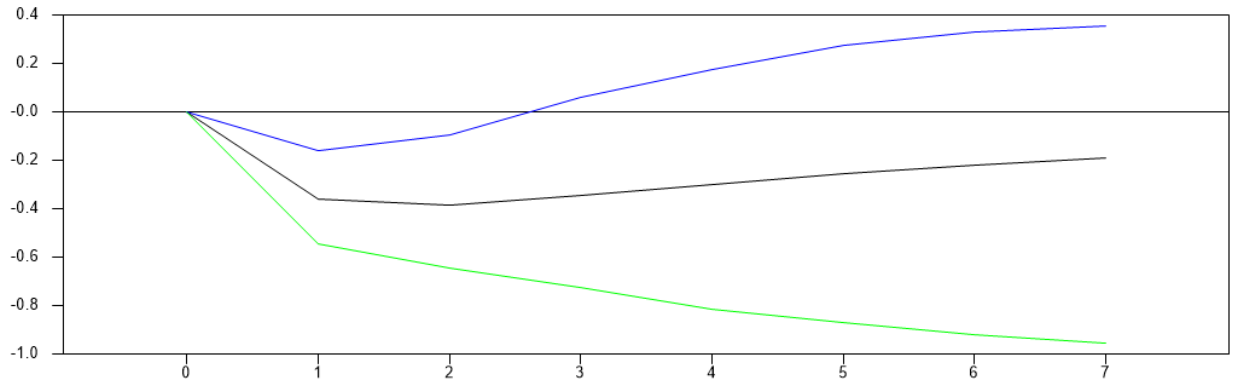
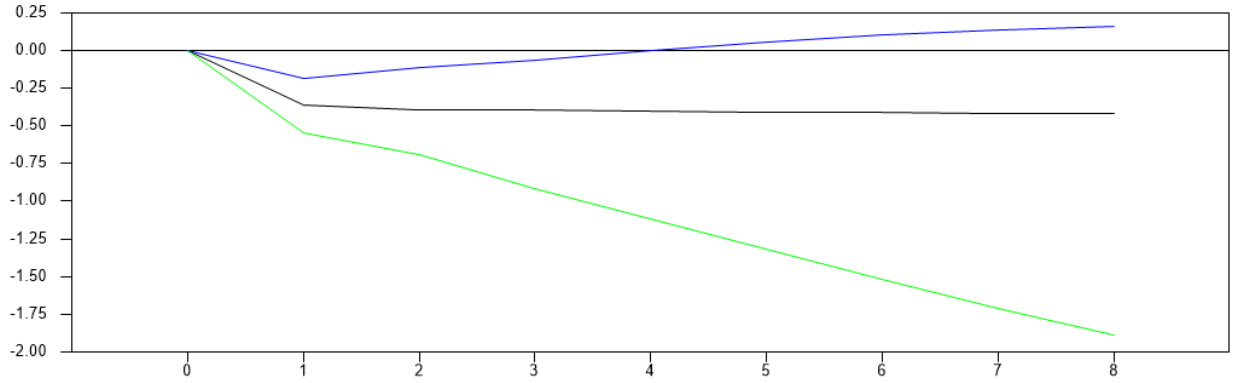


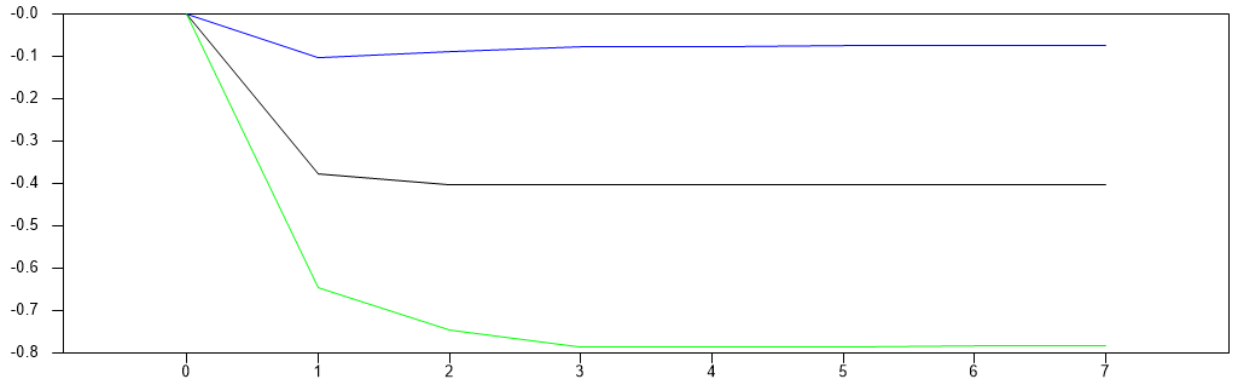
Figure 8: Response of the Level of GDP to a Permanent One Standard Deviation DTF Shock (Baseline Specification, Includes Bootstrapped 10% and 90% Confidence Intervals)



3.4. Growth Rate Specification

Our estimated model assumes both variables are stationary around a trend determined by time effects. It is possible that the correct underlying model for the log of real GDP has a unit root and that perhaps the specification should best be estimated in first-difference form. We performed the panel unit root tests of Levin, Lin and Chu (2002) on both variables and these suggest strong rejections of the null that either series have unit roots. But given the relatively low power of these kinds of tests with small T samples, we decided to estimate the model in first-difference form, thus featuring the growth rates of real GDP per capital and DTF. For comparability, Figure 9 reports the estimated responses to the levels of log GDP per capita to a shock to DTF, obtained by cumulating the impulse responses for the first-difference specification. The magnitude of the estimated decline in GDP is similar to that estimated in the baseline VAR though because of the difference in specifications, it is estimated to be a permanent effect.

Figure 9: Response of the Level of GDP to a One Standard Deviation DTF Shock (Growth Specification, Includes Bootstrapped 10% and 90% Confidence Intervals)



3.5. Arellano-Bond Estimation

Beyond the possibility of unit roots, it is well known that OLS estimates of dynamic panel regressions suffer from biases when T is small. The presence of both country-specific fixed effects and lagged endogenous variables means there are a series of correlations between the error terms and explanatory variables which violate the traditional assumptions required for fixed effect OLS to be an unbiased estimator. As discussed in detail by Bond (2002), there are no easy fixes for these problems, particularly in cases where you have highly persistent dependent variables such as the per capita GDP series used in our paper. One popular method, however, is the Arellano-Bond estimator. This estimator is based on first-differencing the basic specification and then constructing a series of instruments based on the moment conditions which state that that error terms should not be correlated with lagged

levels of the explanatory variable.

Table 6 reports the estimation of our VAR specification using the Arellano-Bond method. These estimates were then used to repeat our Cholesky decomposition method and estimate the impulse reponse of GDP per capital to DTF. To maintain comparability, we estimated the impact of a unit shock to GDP and then scaled the responses using the same standard deviation of DTF shocks reported in Figure 6. As shown in Figure 10, these estimates are relatively similar to our baseline estimates in also showing a short-run negative response, though this effect fades away somewhat quicker and the bootstrapped standard errors do not indicate statistical significance of the responses.

Table 6: Arellano-Bond Estimation of the VAR Model

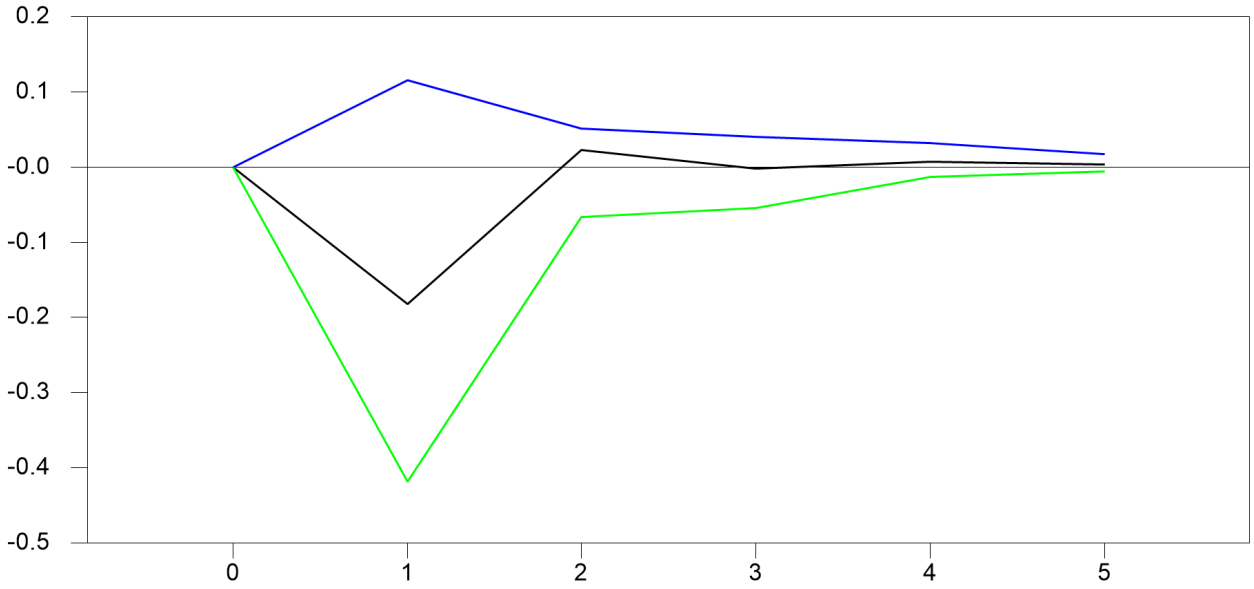
	(1)	(2)
	$\log \text{GDP}_t$	$\log \text{DTF}_t$
GDP_{t-1}	0.247***	0.0217***
	(0.009)	(0.004)
GDP_{t-2}	0.0627***	0.037 ***
	(0.007)	(0.004)
$\log DTF_{t-1}$	-0.063***	0.259 ***
	(0.013)	(0.013)
$\log DTF_{t-2}$	0.040**	0.078***
	(0.013)	(0.014)
N	1971	1971

Notes: Model estimated with time and country dummies.

Standard errors in parentheses

$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Figure 10: Response of the Level of GDP to a One Standard Deviation DTF Shock (Arellano-Bond Estimation, Includes Bootstrapped 10% and 90% Confidence Intervals)



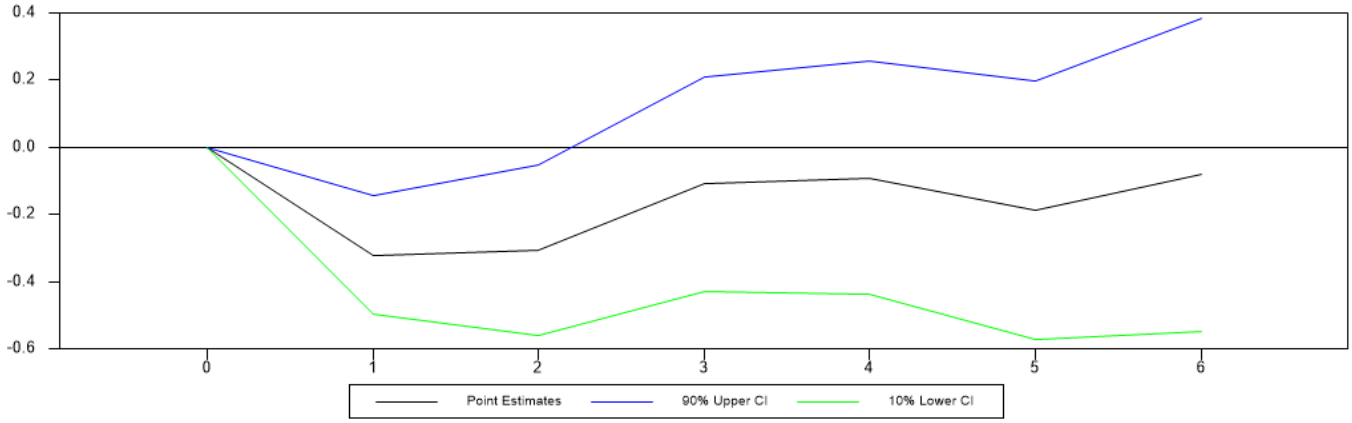
3.6. Local Projection Estimates

The local projection methodology introduced by Jorda (2005) has been commonly used in recent years to estimate impulse responses without having to take a stance on all the parameters governing the dynamics of the variables at hand. We employed this method to estimate impulse responses to a DTF shock as follows. For $j = 1, ..6$, we estimated

$$\log GDP_{i,t+j} = \gamma_i + \mu_t + \beta^j \log DTF_t + \sum_{k=1}^{2} \theta \log DTF_{t-1} + \sum_{k=0}^{2} \theta \log GDP_{t-1} \tag{4}$$

and then use β^j to measure the impulse response to a DTF shock after j periods. The inclusion of contemporanous GDP in the regression specification means, as in our baseline VAR, we are not attributing any contemporaneous effect on GDP of changes in DTF. The estimated standard errors on β^j are used to construct confidence intervals of the same size as the bootstrapped intervals shown earlier. The results are shown in Figure 11. This is not our preferred method for estimating these impulse responses. With a T of 15, we can only use a small fraction of the time dimension of the data set to estimate the longer-dated impulse responses. It is also likely that the standard errors from these regressions are inaccurate due to small sample bias problems. However, the size and pattern of the estimates are similar to those obtained from our baseline VAR.

Figure 11: Response of the Level of GDP to a One Standard Deviation DTF Shock (Local Projection Estimation, Model Includes Two Lags of GDP and DTF, Includes 10% and 90% confidence intervals based on estimated Newey-West standard errors)



3.7. Alternative Measures of Real GDP

One possible explanation is that we are using an inadequate measure of real GDP. It is well known, for example, that movements in the real GDP measures in the WDI do not always correlate well with the movements in the real GDP measures produced by the Penn World Tables.[8] Our measure of GDP does not attempt to make any PPP-related adjustments. Table 7, however, shows that the coefficients from the regression for real GDP in VAR are similar when using two alternative measures: real GDP from the Penn World Tables and the WDI's PPP-adjusted measure of real GDP. The key driver of the negative short-run effect is of a shock to DTF is the negative coefficient on the first lag of DTF and this finding is robust across all three measures. Figures 12 and 13 repeat the VAR impulse response analysis for the impact of DTF shocks on real GDP using the two different measures of real GDP. Both charts back up the conclusions from the previous analysis that shocks to DTF are followed by a negative impact on GDP.

3.8. Using Rank Instead of DTF

We described a number of reasons above why we believe the DTF measure we are using is more appropriate for a panel data analysis the impact of changes in the ease of doing business. The DTF scores are the true indicators underlying the rankings and, as such, one could consider the rankings to be a transformation of the original data that obscure information rather than assist.

Still, one could be suspicious that our specially-constructed DTF time series is somehow responsible for the results that we have found. Figure 14 shows, however, that repeating our analysis using the ease of doing business rank that has been used in other studies, the same results are effectively obtained. A one-standard deviation positive shock to a country's *Doing Business* rank (meaning the country's DTF score is worsening) produces an estimated temporary improvement in real GDP, with the estimated magnitudes being similar to those estimated with the DTF analysis.

[8] See Ram and Ural (2014)

Table 7: VAR Regression for $\log GDP$ with Alternative Measures of GDP

	(1)	(2)	(3)
	$\log \text{GDP}_t$	$\log \text{GDP}_t$	$\log \text{GDP}_t$
	WDI	PPP	PWT
$\log GDP_{t-1}$	0.986***	1.063***	0.979***
	(0.023)	(0.023)	(0.033)
$\log GDP_{t-2}$	-0.060***	-0.155***	-0.177***
	(0.0268)	(0.025)	(0.045)
$\log DTF_{t-1}$	-0.126***	-0.022	-0.220***
	(0.045)	(0.028)	(0.047)
$\log DTF_{t-2}$	0.0114**	0.064**	0.183***
	(0.044)	(0.040)	(0.044)
N	2135	2135	1188

Notes: Model estimated with time and country dummies.

Standard errors in parentheses

$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Figure 12: Impulse Responses of GDP to a One Standard Deviation Shock to DTF Using PPP-Adjusted GDP (Includes Bootstrapped 10% and 90% Confidence Intervals)

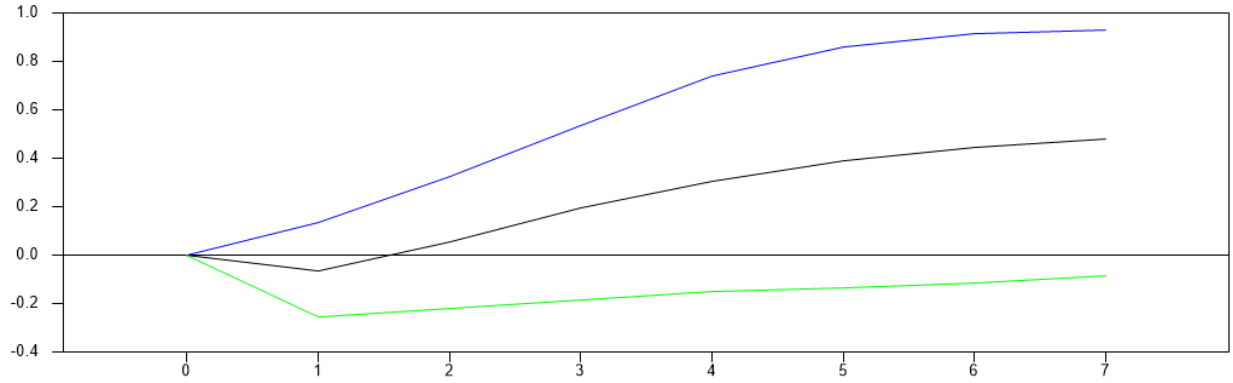


Figure 13: Impulse Responses of GDP to a One Standard Deviation Shock to DTF Using Penn World Tables GDP (Includes Bootstrapped 10% and 90% Confidence Intervals)

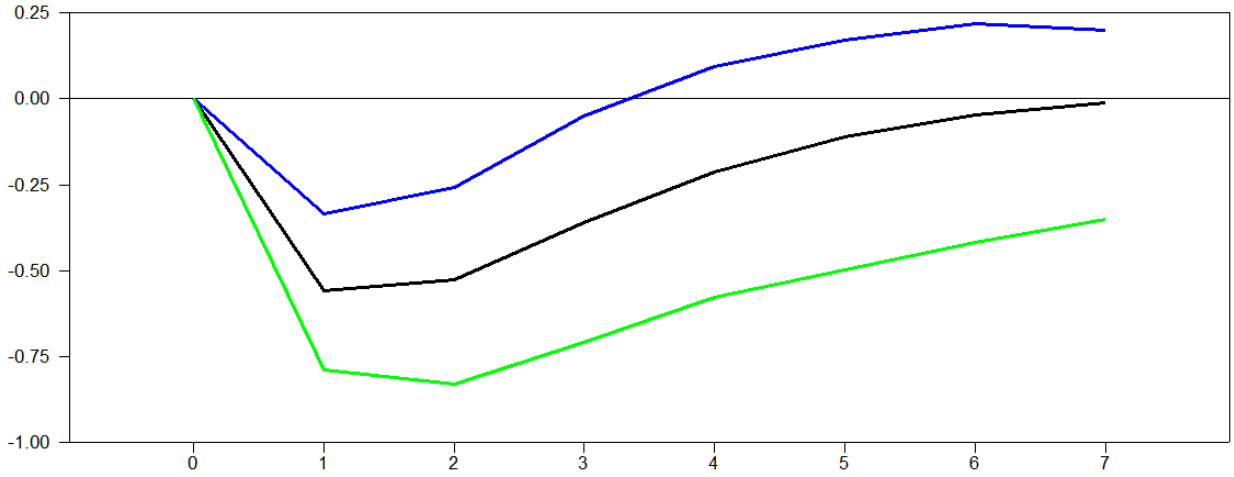
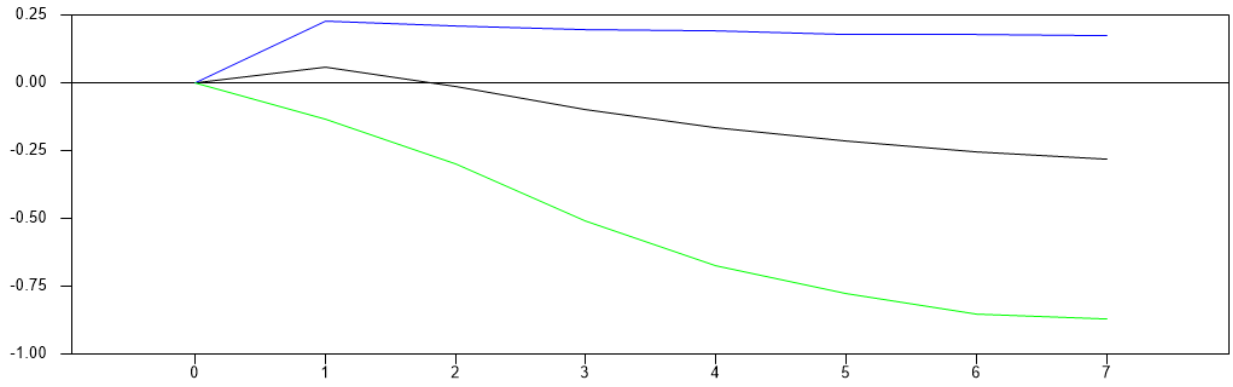


Figure 14: Impulse Responses of GDP to a One Standard Deviation Shock to Doing Business Rank (Includes Bootstrapped 10% and 90% Confidence Intervals)



3.9. Effect on Components of GDP

One question raised by our results is how are they reflected in the reponse of the various sub-components of GDP. One possible explanation for the negative effects of improvements in DTF could be that business-friendly reforms tend to be introduced by conservative governments that are simultaneously cutting public expenditure, so there is a short-run negative impact on GDP due to Keynesian aggregate demand effects. To assess whether this is what is driving the results, Figure 15 shows the results from running our VAR model replacing GDP with expenditure sub-components for consumption, investment and government spending on goods and services. The results do show a point estimate of an immediate reduction in government spending but this is not statistically significant and this is followed by a significant increase in government spending a few years later. The initial response of both consumption and investment is estimated to be negative but roughly in line with the response of GDP, so no clear causal story emerges from looking at these sub-components.

3.10. Alternative Cholesky Decomposition

Recognizing the likely seriousness of the endogeneity problem whereby higher levels of GDP per capita likely improve *Doing Business* scores, we have focused on shocks to DTF that are uncorrelated contemporaneously with shocks to GDP. It is possible that this methodology understates the impact of improvements in the ease of doing business because it discounts potential positive contemporaneous impacts. For completeness, we thus report the results for the alternative identification—with DTF ordered first and the contemporaneous correlation being attributed to DTF immediately impacting GDP—in Figure 16.

As expected, given the contemporaneous correlation of the residuals, the estimated impact of the effect of an improvement in DTF is positive but it is statistically significant. The point estimates for the second and subsequent periods are negative, though not statistically significant. This suggests that even if we can interpret the contemporaneous correlation between GDP and changes in the *Doing Business* as a causal effect running from the DTF score to GDP, the short-run positive impact seems to disappear quickly.

3.11. Using Sub-Components of the DTF Score

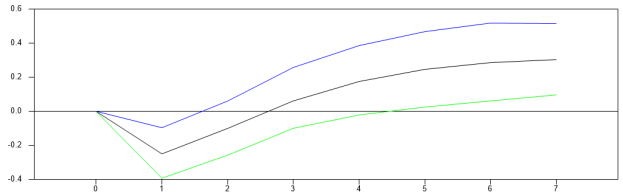
We have also repeated our analysis by replacing the overall DTF score with its various sub-components to assess whether our puzzling findings are driven by particular components of the overall measure. Figure 17 reports that point estimates for impulse responses to improvements in the sub-component indicators are negative over our horizon of seven years for almost all of the indicators. For four of these sub-components, the confidence intervals suggest significant negative effects at various horizons, while the other negative point estimates are not reported as statistically significant.

3.12. Sub-Samples of Different Countries

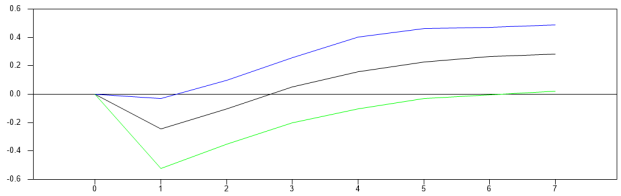
An obvious question is whether our results are driven by a particular sub-set of the countries our sample. One quick way to assess this possibility is to split the sample according to GDP per capita. We have split the sample into quartiles based on the level of initial GDP per capita in the sample. Figure 18 reports the results for the bottom, second, third and top quartiles. The bottom and second quartile show negative initial impacts larger than for the full sample, with larger negative impacts for the bottom quartile. The top quartile shows a small and statistically insignificant positive initial effect followed by significant negative effects. Only the third quartile showed positive, though insignificant, effects.

We also considered other country-related sample splits, such as dividing countries according to their scores on the World Bank's Governance Matters indicators or Transparency International's corruption perceptions index but these splits did not show any systematic patterns.

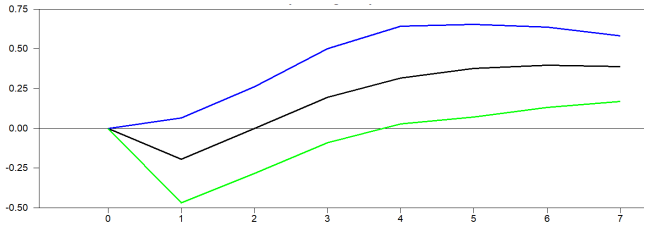
Figure 15: Impulse Responses of components of GDP to a One Standard Deviation Shock to DTF (Includes Bootstrapped 10% and 90% Confidence Intervals)



(a) Effect on Consumption



(b) Effect on Investment



(c) Effect on Government Spending

Figure 16: Impulse Responses of GDP to a One Standard Deviation Shock to DTF Using Alternative Causal Ordering with DTF Ordered First (Includes Bootstrapped 10% and 90% Confidence Intervals)

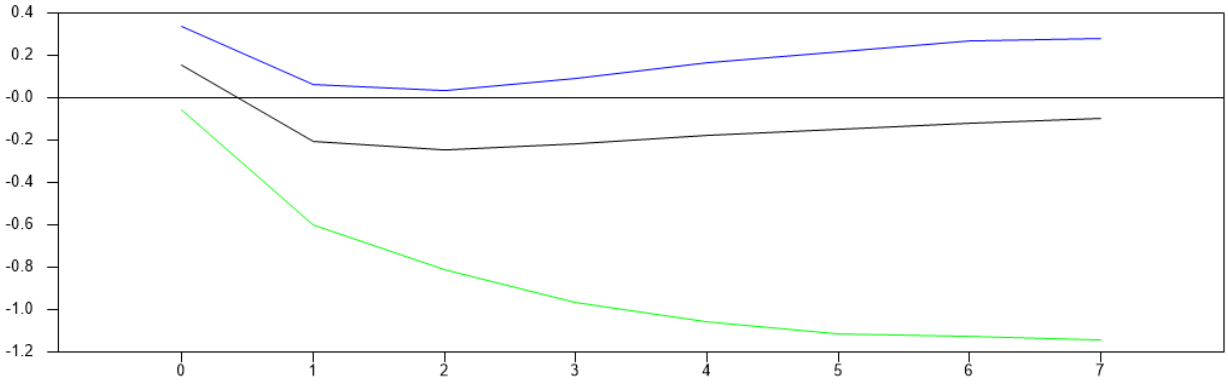
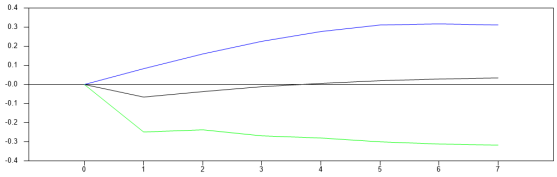
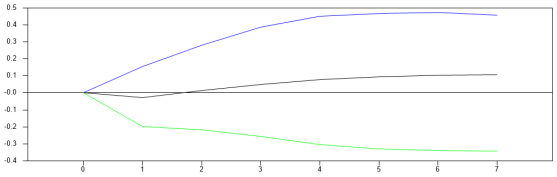


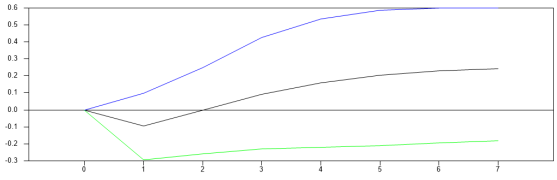
Figure 17: **Impulse Responses of GDP to a One Standard Deviation Shock to Sub-components of Doing Business (Includes Bootstrapped 10% and 90% Confidence Intervals)**



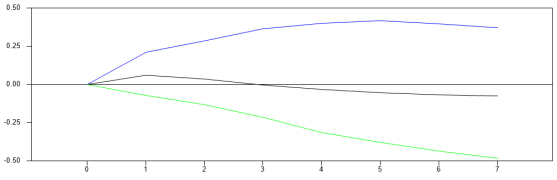
(a) Starting a Business



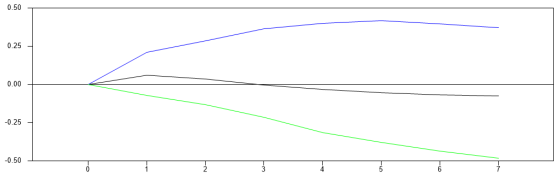
(b) Paying Taxes



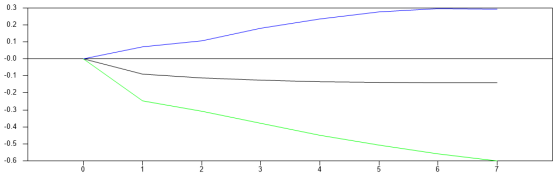
(c) Ease of Trade



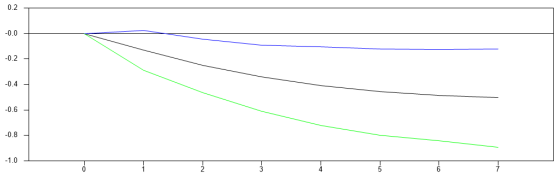
(d) Resolving Insolvency



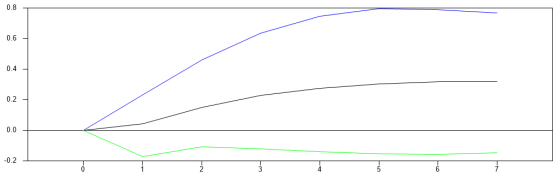
(e) Protecting Minority Investors



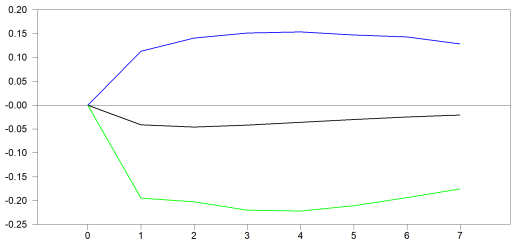
(f) Ease of getting construction permits



(g) Enforcing Contracts

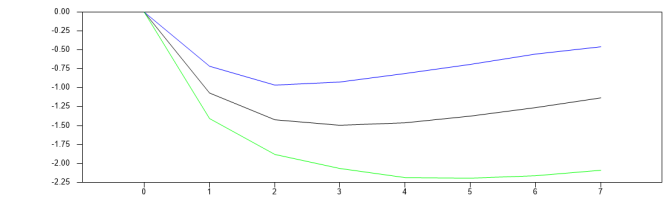


(h) Ease of getting credit

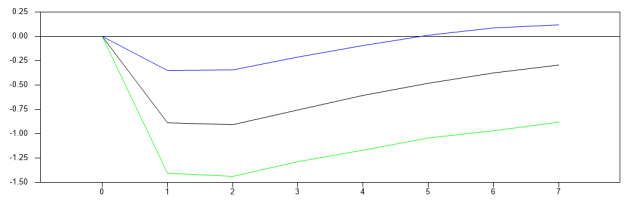


(i) Registering Property

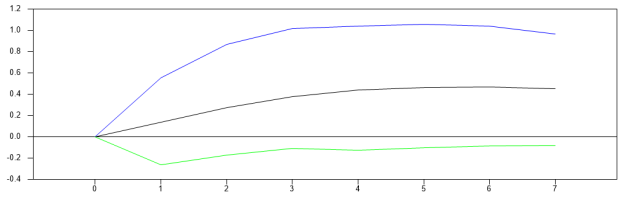
Figure 18: Impulse Responses of GDP to a One Standard Deviation Shock to DTF for four Quartiles of GDP Per Capita (Includes Bootstrapped 10% and 90% Confidence Intervals)



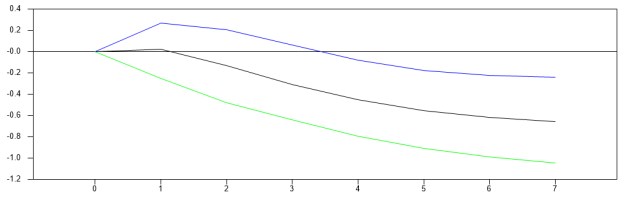
(a) Bottom Quartile



(b) Second Quartile



(c) Third Quartile



(d) Top Quartile

3.13. Sample Split

Finally, one question is whether our results are different for the earlier and later parts of the sample. It may have been that the measures of the ease of doing business became less useful over time as governments focused more on gaming the system. Figures 19 and 20 show the impulse reponse of GDP per capita to DTF for separately estimated samples over 2006-2012 and 2013-2020. For both samples, the point estimates of the response are mainly negative but there is actually more evidence for negative impacts from the early sample than the later sample. One possibility is that there may have been more "low hanging fruit" for governments to focus on in the early years of the *Doing Business* project and these became the focus of reform efforts thus distracting from more substantive policies.

Figure 19: Response of the Level of GDP to a One Standard Deviation DTF Shock (Baseline Specification, Includes Bootstrapped 10% and 90% Confidence Intervals) for sample 2006-2012.

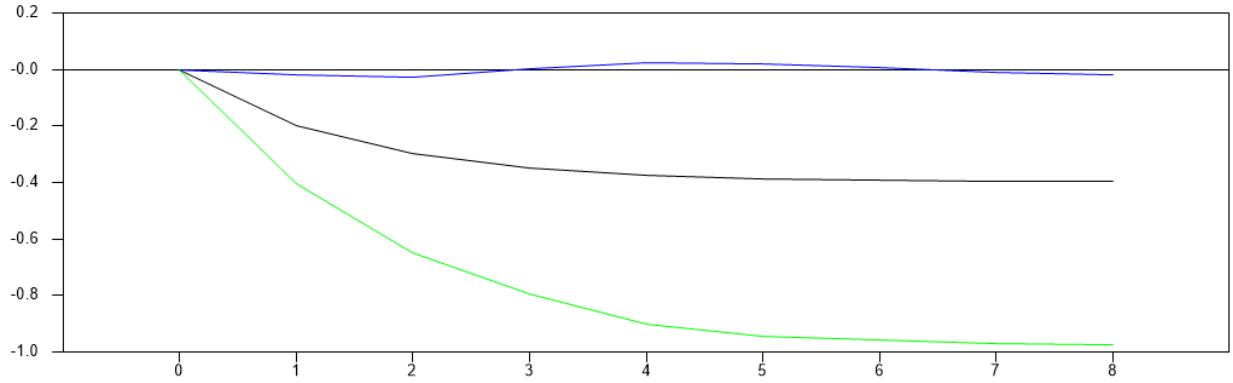
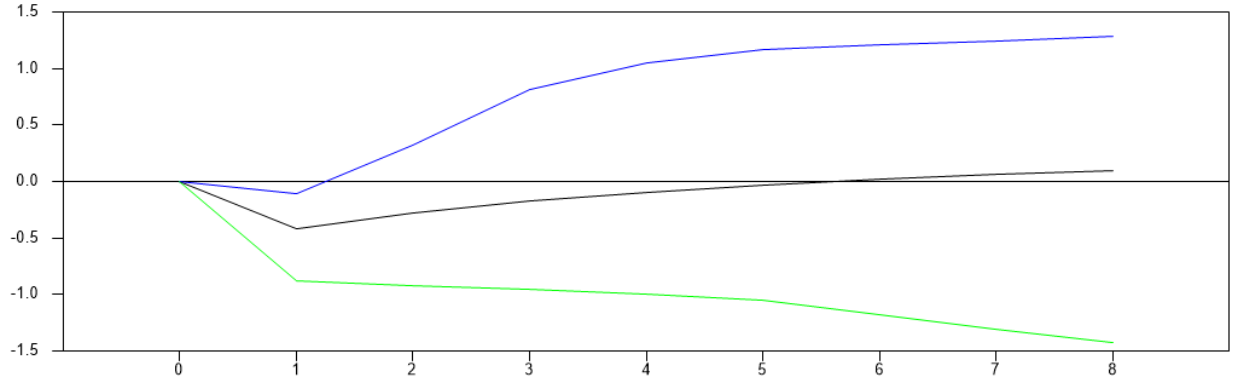


Figure 20: Response of the Level of GDP to a One Standard Deviation DTF Shock (Baseline Specification, Includes Bootstrapped 10% and 90% Confidence Intervals) for sample 2013-2020.



4. Conclusions

In light of previous findings, the results reported here could be considered surprising. However, we believe the methodology of using within-country time series variation to assess the effect of changes in the *Doing Business* score is a valid one and the results turn out to be robust across a wide range of different specifications, data types and estimators. In terms of our failure to find a positive impact of business-friendly reforms on GDP, one possibility is that the *Doing Business* indicators do not do a very good job at measuring the ease of doing business. As noted above, there is a relatively low cross-sectional correlation between the *Doing Business* indicators and comparable estimates from occasional surveys such as the World Bank's *Enterprise Surveys*.

But this still raises an obvious question: Why do we find that improvements in the the *Doing Business* score appear to have had a negative effect on GDP? As discussed above, one possibility is that the widespread focus in the developing world on the *Doing Business* indicators has perhaps had a negative effect, with those countries that have had the best improvements in their DTF scores being countries that have focused on box-ticking exercises to improve their ranking rather than substantive reforms. Another possibility is that implementing improvements in business environment takes take time to have a positive impact—more time than the short time element in our analysis can pick up—and our findings are picking up some shorter-run disruptions that stem from reforms that ultimately have a positive effect. This scenario can be seen in theoretical models with labor market frictions such as Cacciatore, Duval, Fiori and Ghironi (2016). We believe further work could explore the merits of these two explanations.

A final important qualifier to our results is that they clearly can only be considered representative of the sample that we have examined here. It may simply be that the period under investigation in this paper is unrepresentative of the general relationship between the business environment. Certainly, we believe that further investigation of the findings reported here is warranted.

References

[1] Acemoglu, Daron (2008). *Introduction to Modern Economic Growth*, Princeton University Press.

[2] Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* Voulme 91(5), 1369–1401.

[3] Arrunada, Benito (2007). "Pitfalls to Avoid When Measuring Institutions: Is Doing Business Damaging Business?," *Journal of Comparative Economics*, Volume 35, pages 729-747.

[4] Besley, Timothy (2015). "Law, Regulation, and the Business Climate: The Nature and Influence of the World Bank Doing Business Project," *Journal of Economic Perspectives*, Volume 29, Summer, pages 99-120.

[5] Bond, Stephen (2002). "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice," *Portuguese Economic Journal*, Volume 1, pages 141–162.

[6] Cacciatore, Matteo, Romain Duval, Giuseppe Fiori and Fabio Ghironi (2016). *Journal of International Money and Finance*, Volumne 68, page 358-385.

[7] De Long, J. Bradford De Long and Lawrence Summers (1991). "Equipment Investment and Economic Growth," *Quarterly Journal of Economics*, Volume 106, pages 445-502.

[8] Djankov, Simeon (2016). "Correspondence – The Doing Business Project: How It Started," *Journal of Economic Perspectives*, Volume 30, pages 247-248.

[9] Djankov, Simeon, Caralee McLiesh, and Rita Maria Ramalho (2006). "Regulation and Growth," *Economics Letters*, Volume 92, 395-–401.

[10] Freund Caroline and Bineswaree Bolaky (2008). "Trade, Regulations, and Income," *Journal of Development Economics*, Volume 87, pages 309-321.

[11] Gillanders, Robert and Karl Whelan (2014). "Open For Business? Institutions, Business Environment and Economic Development," *Kyklos*, Volume 67(4), pages 535-558.

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

[13] Hallward-Driemeier, Mary and Lant Pritchett (2015). "How Business Is Done in the Developing World: Deals versus Rules," *Journal of Economic Perspectives*, Volume 29, pages 121-140.

[14] Islam, Nazrul (1995). "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, Volume 110, pages 1227-70.

[15] Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian (2013). "Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates." *Journal of Monetary Economics*, Volume 60, pages 255-274.

[16] Jorda, Oscar (2005). "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, Volume 95, pages 161-182.

[17] Kar, Sabyasachi, Lant Pritchett, Spandan Roy and Kunal Sen (2019). Doing Business in a Deals World: The Doubly False Premise of Rules Reform. Effective States and Inclusive Development Research Centre (ESID) working paper, University of Manchester.

[18] Kelly, Morgan (2021). Persistence, Randomization, and Spatial Noise. University College Dublin working paper 21/25.

[19] Lawless, Martina (2013). "Do Complicated Tax Systems Prevent Foreign Direct Investment?," *Economica*, Volume 80, pages 1-22.

[20] Lee, David, Justin McCrary, Marcelo Moreira and Jack R. Porter (2021). Valid t-ratio Inference for IV. NBER Working Paper No. 29124.

[21] Levin, Andrew, Chien-Fu Lin and Chia-Shang James Chu (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, Volume 108, pages 1–24.

[22] Nelson, Richard and Edmund Phelps (1966). "Investment in Humans, Technological Diffusion, and Economic Growth," *American Economic Review*, Volume 56, pages 69-75.

[23] Norback, Pehr-Johan, Lars Persson and Robin Douhan (2014). "Entrepreneurship Policy and Globalization," *Journal of Development Economics*, Volume 110, pages 22-38.

[24] Ram, Rati and Secil Ural (2014). "Comparison of GDP Per Capita Data in Penn World Table and World Development Indicators," *Social Indicators Research*, Volume 116, pages 639-646.

[25] Romer, Paul (1990). "Endogenous Technological Change," *Journal of Political Economy*, Volume 98, No. 5, Part 2, pages S71-S102.

[26] Solow, Robert (1956). "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, Volume 70, pages 65-94.

[27] Staiger, Douglas and James Stock (1997). "Instrumental Variables Regression with Weak Instruments," *Econometrica*, Volume 65, pages 557-586.

[28] World Bank (2017). *Doing Business 2017*.

A Investment Share of GDP

As discussed in Section 3.3, the chart below shows the impulse reponse function for GDP per capita to a one-standard deviation shock to the investment share of GDP, where as with the ease of doing business measure VARs, the investment share is ordered second in the VAR. The responses are positive and statistically significant at all horizons.

Figure 21: Impulse Responses of GDP to a One Standard Deviation Shock to Investment to GDP ratio (Includes Bootstrapped 10% and 90% Confidence Intervals)