



# Working Papers

[www.cesifo.org/wp](http://www.cesifo.org/wp)

## Free Trade Agreements and the Consolidation of Democracy

Xuepeng Liu  
Emanuel Ornelas

CESIFO WORKING PAPER NO. 4103  
CATEGORY 8: TRADE POLICY  
FEBRUARY 2013

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Free Trade Agreements and the Consolidation of Democracy

## Abstract

We study the relationship between participation in free trade agreements (FTAs) and the sustainability of democracy. Our model shows that FTAs can critically reduce the incentive of authoritarian groups to seek power by destroying protectionist rents, thus making democracies last longer. This gives governments in unstable democracies an *extra* motive to form FTAs. Hence, greater democratic instability induces governments to boost their FTA commitments. In a dataset with 116 countries over 1960-2007, we find robust support for these predictions. They help to rationalize the rapid simultaneous growth of regionalism and of worldwide democratization since the late 1980s.

JEL-Code: F130, D720, F530, F150.

Keywords: regionalism, rent destruction, political regimes, trade liberalization.

*Xuepeng Liu*  
*Coles College of Business*  
*Kennesaw State University*  
*2460 Gablewood Dr NE*  
*USA – Marietta, GA 30062*  
*xliu6@kennesaw.edu*

*Emanuel Ornelas*  
*Managerial Economics and Strategy Group*  
*London School of Economics*  
*Houghton Street*  
*UK – London WC2A 2AE*  
*ornelas@lse.ac.uk*

21 January 2013

We thank Facundo Alborno, Pol Antràs, Paulo Arvate, Scott Baier, Richard Baldwin, Jordi Blanes-i-Vidal, Kishore Gawande, Stephen Knack, James Lake, Giovanni Maggi, Edward Mansfield, John McLaren, Stephen Nelson, Lorenzo Rotunno, Alan Spearot, James Raymond Vreeland, Bob Staiger, Ben Zissimos, and seminar participants at several venues for useful comments and suggestions. We also thank Nathan Converse for excellent research assistance. Ornelas gratefully acknowledges research support from the Economic and Social Research Council (grant RES-062-23-1360).

“Striking down trade barriers is critical to sustaining democracy [...] throughout the region.”  
[Former U.S. President George W. Bush at the 2001 summit of the potential signatories of the Free Trade Area  
of the Americas (*New York Times*, 4/18/2001)]

## 1. INTRODUCTION

When the United States formally announced the intention to pursue a free trade agreement with Central America countries, there were three explicit goals, one of which was “to support democracy in the region” ([www.whitehouse.gov](http://www.whitehouse.gov), 16 January, 2002). Indeed, the establishment of new democracies has often been followed by the formation of preferential arrangements (or the accession to existing ones). This was the case, for example, of all Mercosur members, of Greece, Portugal and Spain in their accessions to the European Community, and of the EU agreements with Central and Eastern Europe countries shortly after the fall of the iron curtain. It is therefore not too surprising that governments regularly report to the World Trade Organization that “promoting democracy and political stability” is a central force behind their decisions to form regional trade agreements (World Trade Organization 2011). Of course, this may be mere rhetoric. But maybe not.

In this paper we develop a coherent theoretical mechanism for the link between participation in free trade agreements (FTAs) and democratic consolidation. We then scrutinize it empirically. Specifically, we argue that participation in FTAs can serve as a commitment device to destroy future protectionist rents. Since such rents are attractive to autocratic groups, FTAs lower their incentives to seek power. While this may have little value for established democracies, where the rule of law is strong and the risk of authoritarian disruption is negligible, it matters for unstable democracies. Some states will therefore have an *extra* incentive to seek involvement in FTAs, over and above the agreements’ potential trade gains.

We provide the theoretical basis for our claims by extending the preferential trade integration model developed by Ornelas (2005a) to allow for endogenous changes in the political regime.<sup>1</sup> In that otherwise standard model, under any trade regime domestic firms exchange transfers for protection with the government, which cares about national welfare and the transfers it receives. The government then decides whether to form an FTA considering the political economy equilibrium under each trade regime. The key to understand the impact of an FTA is the recognition that the equilibrium of the (*ex post*) external tariff game changes under the constraint imposed by the agreement on the internal tariffs. Once one takes this into account, and only then, one finds that even

---

<sup>1</sup> To our knowledge, McLaren (2000) provides the only other theoretical analysis of how political turnover affects governments’ decisions to participate in trade agreements. However, his emphasis is very different, on the choice between free trade areas and customs unions.

though FTAs still permit lobbying for protection against excluded countries, the volume of protectionist rents falls with the formation of an agreement.

In a dynamic setting this implies that, all else equal, groups motivated mainly by rents will have lower incentives to seek power if the country is deeply engaged in FTAs and withdrawal from the agreement is costly. Authoritarian groups tend to fit this description well, given their aptitude to resort to violence rather than to rely on accountability to keep power.<sup>2</sup> If the gain of authoritarian groups from seeking power falls when the country is engaged in FTAs, but the agreement does not alter the costs and risks from attempting a coup d'état, the likelihood of democratic failure will, all else being equal, be lower if the country is more intensively involved in FTAs. Hence, greater participation in FTAs increases the likelihood of democracy survival.

If the incumbent government in an unstable democracy realizes this effect of “democratic consolidation,” it will seek participation in FTAs more actively than it otherwise would, in order to weaken the authoritarian threat. Yet even if the dictatorial group takes control despite the FTA, the agreement would still constrain its rent-extraction activities. For both of these reasons, unstable democracies tend to enter in FTAs more frequently than stable ones.

Analyzing the formation of FTAs and the strength of democracy in 116 countries over 1960-2007, we obtain empirical support for both of our main theoretical results. Employing duration analysis techniques, we find that greater participation in FTAs lowers the likelihood of democracy failure in a country. Using the estimated hazard rates from the duration analysis, but considering only the portion of the likelihood that is not predicted by current FTA participation, we find as well that a higher risk of democratic breakdown induces countries to participate more actively in FTAs. Our empirical results are robust to many different econometric specifications, alternative measures of democratic transition and different sets of controls.

One of our empirical challenges is to define *how* unstable a democracy is. We do so by relying on the approach proposed by Persson and Tabellini (2009), who estimate the likelihood of democracy failure employing the concept of “democratic capital.” The domestic component of democratic capital takes into account the history of democracy in the country. The longer the country has experienced democracy, and the more recent is its democratic experience, the greater the country's stock of domestic democratic capital. The foreign component of democratic capital encompasses instead current levels of democracy abroad. The greater the number of democratic

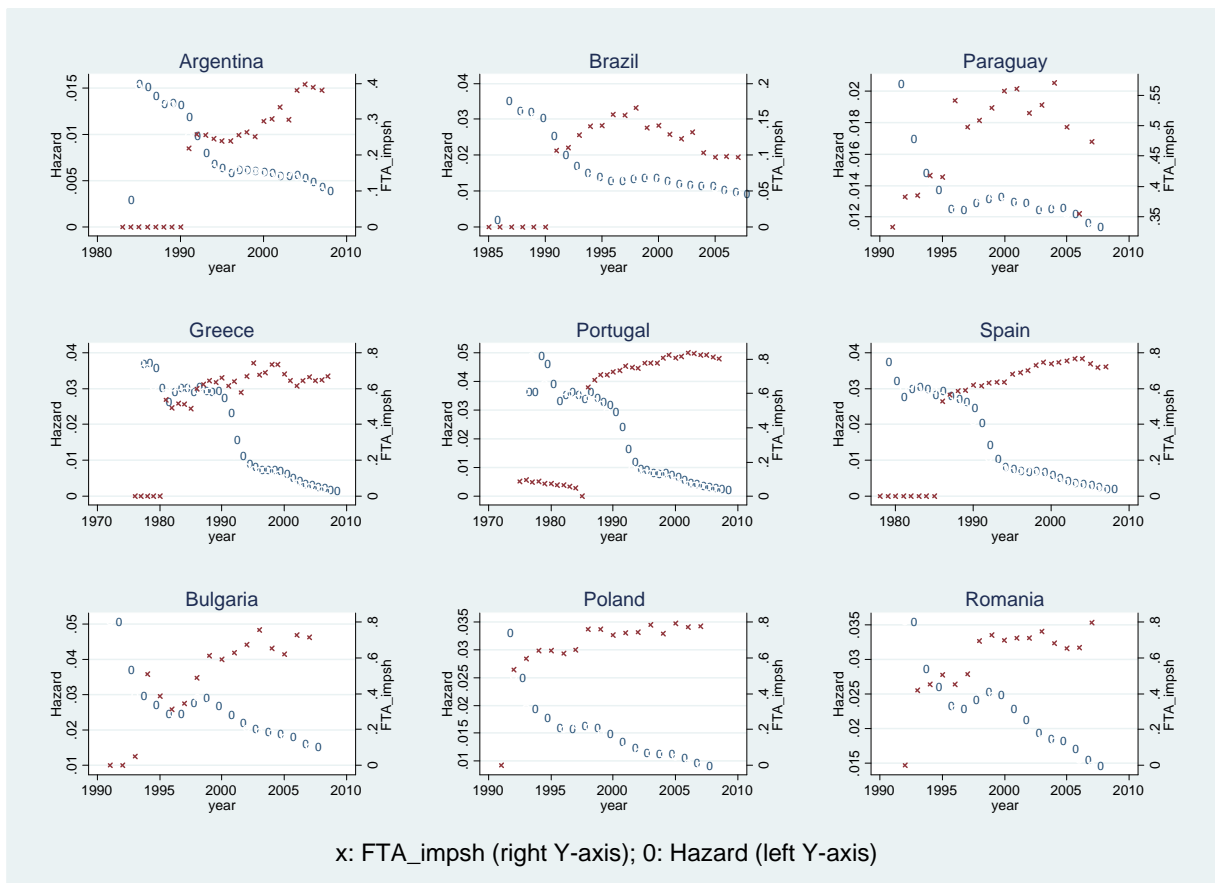
---

<sup>2</sup> In our analysis we do not take a stance on whether democratic or authoritarian regimes are more rent-seeking. We nevertheless note that, at least in the trade context, authoritarian regimes are often associated with greater restrictions (see e.g. Aidt and Gassebner 2010, Mansfield et al. 2000 and Mitra et al. 2002).

countries, and the closer those democracies are to a country, the greater is that country's stock of foreign democratic capital. Along with other covariates, these two components of democratic capital allow us to estimate the likelihood of democracy failure in a country. We find that greater participation in FTAs reduces this probability.

Our estimates are statistically significant and also economically meaningful. Figure 1 illustrates this. It plots, for several countries, the estimated hazard against the share of imports that stem from FTA partners.<sup>3</sup> As the figure clearly shows, the hazard out of democracy drops sharply right after those countries become significantly engaged in FTAs.

Figure 1: Hazard out of democracy and share of imports from FTA partners, selected countries



Notes: "FTA\_impsh" denotes the share of the country's imports that originates from FTA partners. "Hazard" is the hazard out of democracy, estimated as explained in section 4.

A concern is that FTA participation may be endogenous in the estimation of democracy

<sup>3</sup> It is worth noting that the relationship is not driven by the countries illustrated in Figure 1; in fact, we show that our finding holds well even after we drop Mercosur and Eastern European countries from the analysis.

longevity. Indeed, our second hypothesis implies a positive relationship between democracy instability and formation of new FTAs. From this view the endogeneity bias would imply a *positive* relationship between FTA participation and democracy instability, suggesting that the negative relationship we obtain is a lower bound (in absolute value) for the true effect. We nevertheless investigate further this issue and propose an instrumental variable estimation that relies on “contagion” effects in FTA formation that follows a reasoning advanced by Egger and Larch (2008) and Baldwin and Jaimovich (2012). The main idea is that, when its neighbors engage more deeply in FTAs, a country would tend to follow suit. We confirm that contagion measures are indeed strong instruments for FTA participation. Moreover, our finding that more intense FTA participation increases the longevity of democracies is not altered when we use this approach.

Interestingly, our predictions hold consistently only for full-fledged agreements, signed under GATT’s Article XXIV. Article XXIV requires free trade agreements to cover most of the trade among the members and that members liberalize fully vis-à-vis each other. By contrast, agreements signed under the Enabling Clause of the GATT permit many exceptions and are often not fully implemented.<sup>4</sup> We do not find any significant association between those partial-scope preferential trade agreements (as well as those signed but not reported to the WTO) and democracy survival, or between political instability and the formation of partial-scope agreements. These stark differential results are consistent with the rent-destructing mechanism we put forward, as the partial-scope arrangements, unlike those signed under Article XXIV, impose very few restrictions on the availability of rents from protection.

To our knowledge, Pevehouse (2005) is the only other to relate participation in trade agreements—in fact, participation in international organizations broadly defined—to the durability of democracy. He argues that joining (although not the participation itself) an international organization tends to increase the longevity of new democracies, mostly because of signaling effects, and provided that the international organization has high “democratic density.” Following Pevehouse, we also test whether FTAs help to sustain democracies only when the arrangements are composed mostly of highly democratic members. We find that the role of an FTA in helping to sustain a country’s democracy is not greater when the agreement is with more democratic partners. Thus, pressure from more democratic FTA partners is not what drives our results.

Our second finding is also novel, but relates to other previous results. Mansfield, Milner and Rosendorff (2002) find that pairs of democratic countries are more likely to share a trade agreement

---

<sup>4</sup> This relates to the findings of Mrazova et al. (2012), who show that Article XXIV can have important welfare implications.

than pairs in which at least one country has an authoritarian political regime.<sup>5</sup> We take the Mansfield et al.'s finding on board and focus on democracies. We then dig deeper to understand whether democracies at different stages in their "consolidation" process have different incentives to engage in FTAs. We find that they do: among democracies, those who face threats to their democratic regimes are especially prone to form FTAs.

Mansfield and Pevehouse (2006), like us, distinguish between different types of democracies and their willingness to join international organizations. Specifically, they find that countries that have undergone a transition to democracy in the previous five years are more likely to join international organizations, in particular those made up of more democratic states. Two central differences between the findings of Mansfield and Pevehouse and our second main result are worth emphasizing. First, we explicitly estimate each country's hazard out of democracy (rather than splitting democracies between new and old ones). Second, we do not pool all types of international organizations together. This difference matters, as we find a strong effect of democratic instability on FTA participation but not on participation in partial-scope trade agreements.

These (and other) differential effects do not prove, but are all consistent with the idea that engagement in FTAs helps democracies to endure because of their rent destruction effects. Our novel measure of FTA participation further reinforces the view that the destruction of rents is an essential element behind the result. Rather than using dummies to indicate participation in FTAs, we use the share of a country's imports that come from its FTAs partners. As we argue in section 4, this continuous measure of FTA participation is far superior to the common practice of using dummies. Furthermore, it is supported by the model, where we show that our measure is negatively correlated with the volume of available protectionist rents under FTAs.

The essence of our argument is related to the commitment rationale for trade agreements espoused by Maggi and Rodriguez-Clare (1998, 2007), although in their analyses a government enters a free trade agreement not to affect its successor's policies, but its own (otherwise time-inconsistent) future policies. The more general idea that governments can manipulate state variables to constrain their successors' choices was first advanced in the macroeconomics political economy literature.<sup>6</sup> More recently, Acemoglu and Robinson (2006) have developed a general framework to study circumstances when an incumbent democratic government can design economic policies to

---

<sup>5</sup> Mansfield, Milner and Pevehouse (2008) find that this holds for different types of trade agreements except the "shallowest" ones, according to a five-tier classification. Roy (2011) takes a distinct perspective, showing that the breadth of the service commitments undertaken by WTO members is greater for more democratic countries.

<sup>6</sup> See e.g. the pioneering contributions by Alesina and Tabellini (1990) and Persson and Svensson (1989).

irreversibly change the expected net benefit of future coups.<sup>7</sup> A related reasoning is employed here to show that a democratic government, when faced with the prospect of political disruption, may want to limit the ability of a potential authoritarian government to create rents through trade policies. We innovate in this dimension by showing, theoretically and empirically, that an FTA can be an effective tool for that purpose.<sup>8</sup>

There is also an important line of research that links democracy to trade liberalization and openness.<sup>9</sup> The forces typically emphasized in that literature are however quite different from the mechanism we advance here. Moreover, and critically, in this paper we focus on the role of trade *agreements*, where an external commitment makes the policy costly enough to reverse so that it can credibly affect the actions of future governments.<sup>10</sup> A unilateral tariff reduction would clearly not fulfill this requirement.

The paper proceeds as follows. In section 2 we describe the model. In section 3 we analyze the incentives to form a free trade agreement. We discuss our empirical strategy in section 4. The data are described in Section 5. We show our empirical results in section 6. We conclude in section 7.

## 2. THE MODEL

### 2.A. The economic structure

We consider a 3-country,  $N$ -sector competitive economy where in each sector there is a “natural importer” country that would import the good from the other two countries under free trade. Goods are produced under constant returns to scale. One unit of the numeraire good 0 is produced with one unit of labor. All other goods  $j = 1 \dots N - 1$  are produced with labor and a sector-specific factor. Thus, whenever good 0 is produced in equilibrium, which we assume to be the case, the wage rate equals unity and general equilibrium forces are absorbed by that sector.

The analysis is conducted from the perspective of a “Home” country. Home’s population consists of a continuum of agents with measure one. Each agent is endowed with one unit of labor,

---

<sup>7</sup> They demonstrate, for example, how trade and capital account liberalization reduce equilibrium taxation under democracy while also rendering coups more costly through the impact of openness on factor prices.

<sup>8</sup> This contrasts with an alternative (and rather ubiquitous) view in the trade literature that regards FTAs as *rent-creating* devices (e.g. Grossman and Helpman 1995). See Freund and Ornelas (2010) for a discussion of that literature.

<sup>9</sup> See for example Giavazzi and Tabellini (2005), O’Rourke and Taylor (2007), Lopez-Cordova and Meissner (2008), and Stroup and Zissimos (2012). There is, of course, also a vast literature on the determinants of democracy and of its durability. See Barro (1999) for a classic reference for the former and Przeworski et al. (1996) for the latter. Gassebner et al. (2012) provide empirical scrutiny of many factors that could affect the survival of democracies.

<sup>10</sup> See Bagwell and Staiger (2002, 2010) for a discussion of this and other roles of trade agreements.



whereas specific factors are owned by a negligible fraction of the population. Consumers have quasi-linear utility of the form  $U = q^0 + \sum_{j=1}^{N-1} [Aq^j - (q^j)^2 / 2]$ , which generates demand  $D^j = A - p^j$  for good  $j$ .

Home is the natural importer of goods  $m = 1 \dots M$ , country  $Y$  is the natural importer of a subset  $E$  of different goods, and country  $Z$  is the natural importer of the remaining  $(N - M - E - 1)$  non-numeraire products. Home's owners of the specific factor used in sector  $j$  earn  $\pi^j(p^j)$ , where  $p^j$  denotes the price of good  $j$  in Home's market. In the non-numeraire sectors, the domestic supply of each imported good  $m$  is  $S^m(p^m) = d^m p^m$  and the supply of each exported good  $x$  is  $S^x(p^x) = d^x p^x$ , where  $d^x > d^m > 0$ . An analogous specification applies for the supply and demand structures of countries  $Y$  and  $Z$ . Home can use specific import tariffs in each import sector; other policy instruments are assumed unavailable. We represent Home's tariff on imports from country  $j$  by  $t_j$ ,  $j = Y, Z$ . Because all import sectors are identical, we will write prices and tariffs without sector-identifying superscripts.

Prices in the three countries are linked by arbitrage conditions. For a generic product imported by Home, this condition is

$$(1) \quad p = p_Y + t_Y = p_Z + t_Z,$$

provided that tariffs are not prohibitive. Using this arbitrage condition, market-clearing requires

$$(2) \quad D(p) - S^m(p) = S^x(p - t_Y) - D(p - t_Y) + S^x(p - t_Z) - D(p - t_Z).$$

Using the expressions for demand and supplies defined above, condition (2) can be rewritten as

$$(3) \quad \hat{p}(t_Z, t_Y) = \gamma + (t_Z + t_Y)\rho,$$

where  $\gamma \equiv 3A / (3 + d^m + 2d^x)$  and  $\rho \equiv (1 + d^x) / (3 + d^m + 2d^x)$ .

When Home is not a member of a free trade agreement, it follows GATT's requirement of non-discrimination. When Home is in an FTA, imports from the FTA partner are duty free, but imports from the excluded country remain taxed, although the country's optimal external tariff will in general change as a result of the FTA.

## 2.B. The political structure

We consider that any group represented in the government enjoys power because there are rents for holding office. The sources of those rents are transfers/bribes, which the private sector offers to government officials in exchange for more favorable policies. Thus, the rents are specific to incumbency, as in models like Besley and Coate's (2001).

Policymakers also care about national welfare. Numerous reasons can explain this concern.

For example, the policymaker could represent a large group in the society, which benefits from a more prosperous economy. Or maybe policymakers that are good at promoting national welfare obtain more public support, which may affect the duration of their office tenure. Since modeling the precise way in which policymakers form their preferences is not essential for our analysis, we take an agnostic view and simply assume that whoever is in power sets policy considering both its welfare consequences and its capacity to attract transfers.

Let us define the measures of welfare. Welfare generated in a specific-factor import sector is denoted by  $W^m(t)$ , whereas  $W^x$  represents welfare from a specific-factor export sector. The former is defined as the sum of consumers' surplus, tariff revenue and producers' surplus generated in that sector; the latter is defined as the sum of consumers' and producers' surplus in the sector. Welfare aggregated across all non-numeraire import and export sectors is then  $W^M(t) \equiv MW^m(t)$  and  $W^X \equiv (N - M - 1)W^x$ , respectively.<sup>11</sup> National welfare,  $W(t)$ , aggregates welfare across all sectors:

$$W(t) \equiv 1 + W^M(t) + W^X = 1 + \sum_{m=1}^M W^m(t) + \sum_{x=M+1}^{N-1} W^x.$$

The preference of the government is specified as

$$(4) \quad G(t, T) \equiv \sum_{m=1}^M G^m(t, T^m) + \sum_{x=M+1}^{N-1} G^x,$$

with  $G^x \equiv W^x$  and

$$(4') \quad G^m(t, T^m) \equiv W^m(t) + bT^m,$$

where  $T^m$  denotes the transfer from import-competing sector  $m$  to the government,  $T \equiv \sum_{m=1}^M T^m$ , and  $b > 0$  reflects the "rent-seeking bias" of the government, or how susceptible to bribes/transfers its policies are. Thus, if for example the government's rent-seeking bias were very high, it would care mainly about rents. Notice that the government sets policy according to (4) regardless of its nature, democratic or not. We adopt this assumption not because we believe that both types of governments implement identical policies. It is simply that our main results do not depend on this distinction, and we have nothing to add to the (extensive) debate on whether democracies or autocracies are more rent-seeking.<sup>12</sup> We also abstract from domestic political competition. Nevertheless, one can reinterpret the government's payoff as an expected payoff (taking into account the probability of

---

<sup>11</sup> Note that we denote welfare in import-competing sectors as a function of the tariff, but not in export sectors. In reality,  $W^x$  also depends on tariffs, but on those imposed by foreign countries  $Y$  and  $Z$ . Since those tariffs are given from the perspective of the Home government under any trade regime, we employ this more concise representation for notational ease.

<sup>12</sup> Grossman and Helpman (1996) provide microfoundations for the weights in equation (4'), but in a model of electoral competition, whereas our context is one where a potential autocrat considers taking over the country, not through the ballot box but through force. Dixit (2010) provides a nuanced general discussion of rent-seeking behavior in democracies and autocracies.

being voted out of power). Ultimately, what is essential is that the rents from holding office depend on the policies implemented.

We assume that producers within each industry can overcome free-riding problems in their lobbying activities. Because of the symmetry and independence across sectors, we focus on a single import-competing sector. The net payoff of producers in such a sector corresponds to the industry's aggregate profits,  $\pi^m(t)$ , subtracted of the transfers it gives to the local government,  $T^m$ .

As in Maggi and Rodríguez-Clare (1998), we model the interaction between government and each domestic industry as a Nash bargaining game, where each side obtains half of the total surplus from the negotiation process.<sup>13</sup> Under the Nash bargaining protocol, the outcome of the bargaining process is jointly efficient. Thus, the “political tariff” resulting from this interaction satisfies

$$(5) \quad t^p = \arg \max [W^m(t) + b\pi^m(t)],$$

where the term in brackets represents (up to a constant) the joint payoff of the government and the industry in a representative import-competing sector. We concentrate on the case where the solution to problem (5) is interior. This corresponds to assuming that  $b < b_{max} \equiv (1+d^m)(d^x-d^m)/(1+d^x)d^m$ .

## 2.C. Equilibrium payoffs

If the private sector were able to capture the entire surplus from lobbying, it would only need to compensate the government for the distortions that  $t^p$  creates. In this case, the government would obtain just its reservation payoff, which is equivalent to how much it could get in the absence of lobbying activities, when it could do no better than by setting the tariff in each import sector to maximize national welfare, yielding a payoff from each import-competing sector of  $W^m(t^p(b=0))$ , or simply  $W^m(b=0)$ . This would require a transfer from each lobby of  $[W^m(b=0) - W^m(b)]/b$ , where  $W^m(b) \equiv W^m(t^p(b))$ . Conversely, if the government could retain the whole surplus from lobbying, producers from each import-competing industry  $m$  would earn only their reservation payoff of  $\pi^m(b=0) \equiv \pi^m(t^p(b=0))$ , entailing a transfer of  $[\pi^m(b) - \pi^m(b=0)]$  to the government. Since government and domestic industry split the surplus from lobbying, it is easy to see that the equilibrium transfer from each industry to the government is  $T^m = [\pi^m(b) - \pi^m(b=0)]/2 + [W^m(b=0) - W^m(b)]/2b$ . Hence, in equilibrium the government obtains from each non-numeraire import-competing sector a payoff of

---

<sup>13</sup> One may want to distinguish the bargaining power of the government relative to the domestic industry depending on whether it is democratic or autocratic. For example, one may argue that the forces limiting rent-seeking behavior are weaker in a dictatorship because autocracies are less accountable, implying a higher bargaining power for the government in autocracies. Since this has no bearing on our results, we take the simpler route of assuming that government and private sector always split the surplus from lobbying.

$$(6) \quad G^m = W^m(b) + bT^m(b) = W^m(b) + b \left\{ \frac{1}{2} [\pi^m(b) - \pi^m(b=0)] + \frac{1}{2b} [W^m(b=0) - W^m(b)] \right\}.$$

There is a more intuitive way of representing this expression. First, define the “political rents” created in the lobbying process in each non-numeraire import-competing sector as

$$(7) \quad PR^m \equiv [W^m(b) + b\pi^m(b)] - [W^m(b=0) + b\pi^m(b=0)].$$

The expression in the first brackets of (7) is the maximized joint payoff of the government and the industry, while the expression in the second brackets is the value of the same function in the absence of lobbying. Thus, the difference between them represents the surplus that the lobbying process adds to the joint payoff of government and industry. Using expressions (6) and (7) we can see that, in equilibrium, the government obtains its reservation payoff in the sector plus its share of the political rents:

$$(8) \quad G^m = W^m(b=0) + \frac{PR^m}{2}.$$

Aggregating across all sectors, we can write (4) evaluated at the equilibrium as

$$(9) \quad G \equiv G(t^p, T) = W^M(b=0) + W^X + \frac{PR}{2},$$

where  $PR \equiv \sum_{m=1}^M PR^m$ . Hence, the government obtains in equilibrium its reservation utility,  $[W^M(b=0) + W^X]$ , plus its share of the political rents. This makes clear that the group in power does not fully internalize the welfare distortions due to its use of the political tariff.

In contrast, if the same political group were *out* of power, its payoff would be different even if the tariff were the same. The reason is that the group does not receive any rents if it is not in a position to enact policies. Accordingly, in that case the group would receive none of the available office rents, and its equilibrium payoff  $H$  would reflect only the general state of the economy:

$$(10) \quad H \equiv G(t^p, 0) = W^M(b) + W^X.$$

Since  $W^M(b=0) \geq W^M(b)$  and  $PR \geq 0$ , it follows directly from (9) and (10) that there are benefits from holding office.

## 2.D. Coup threat

We consider a simple 2-period environment. In the first period there is a democratically elected government in power. There is however a group of citizens representing a segment of the population that is not in the government, which may attempt to take power through force. There is an exogenous probability that they will have this opportunity. We are agnostic about the identity of the citizens represented by this group; it could be the military as well as part of the country's

capitalists or the upper class, for example. In any case, if a coup is attempted and is successful, the authoritarian group takes power in the second period.

Naturally, numerous factors affect both the possibility of initiating a coup and its probability of success. For example, both tend to be higher the stronger the “support” of the pro-coup citizens and the weaker the “resistance” from the segments of the population opposed to the coup. We do not model explicitly these probabilities, because quantifying those forces would be remarkably difficult. We highlight, however, that they are likely to be strongly affected (negatively) by the country’s stock of “democratic capital” (*DC*). The notion of democratic capital, introduced by Persson and Tabellini (2009), proxies the strength of the country’s democratic institutions, and in this sense it provides a useful and concise way of capturing several forces highlighted in the voluminous literature on the durability of democracies. Specifically, in the definition of Persson and Tabellini (2009) the current stock of *DC* in a country is determined by both the level of democracy in the country’s neighbors and by the country’s democratic history. Accordingly, they reason that in nations with enduring democratic tradition, where the rule of law is strong, democratic capital will be abundant and significantly limit the possibility of political disruption. Conversely, in countries lacking solid institutions, where the rule of law is weak, democratic capital will be scarce, thus opening a tangible opportunity for successful coups. Since the level of democratic capital in a country can be considered exogenous (or at least pre-determined) to the relevant political groups, we will rely on it in our empirical analysis.

Both the democratic government and the authoritarian group discount future payoffs according to a common discount factor  $\delta \in [0, 1]$ . To understand when the authoritarian group will attempt to subvert the country’s democratic order, we model the group’s problem as simply as possible. In particular, we assume that, if the takeover attempt is successful, the authoritarian group imposes an autocratic regime in the country and obtains its office payoff  $G$  in the second period. If the takeover attempt is unsuccessful, the group bears a fixed cost  $K > 0$ .<sup>14</sup>

When a coup is attempted, the present value payoffs of the incumbent and of the authoritarian group are represented, respectively, as follows:

$$(12) \quad \Gamma^D = G + \delta[(1 - \rho)G + \rho H] ,$$

$$(13) \quad \Gamma^A = H + \delta[(1 - \rho)(H - K) + \rho G].$$

If no coup were attempted, the incumbent government and the authoritarian group receive,

---

<sup>14</sup> Parameter  $K$  provides a proxy for the many kinds of penalties that could apply in such a case—incarceration, extradition, death and the like.

respectively,  $G$  and  $H$  in each period.

The (risk-neutral) authoritarian group attempts to take power if and only if the expected utility from the endeavor is positive:  $\Gamma^A > (1 + \delta)H$ . Using (13), this condition is equivalent to

$$(14) \quad \rho(G - H) > (1 - \rho)K,$$

where  $\rho$  denotes the probability of success of the coup. That is, provided that it can initiate a coup, the authoritarian group will attempt to take power if its expected gain from seeking power is large relative to the expected cost of a failed coup. To make explicit what is behind this decision, we use expressions (9) and (10) to rewrite condition (14) as

$$(15) \quad [W^M(b=0) - W^M(b)]_+ \frac{PR}{2} > \frac{1 - \rho(DC)}{\rho(DC)} K.$$

In a consolidated democracy, where democratic capital is very high, an attempt against the country's democratic system is unlikely, unless the costs of failure are too low—which is rarely the case—or the gains from holding power are very significant. Our central goal is to analyze how a free trade agreement affects the latter, and through that channel the endurance of democracy in a country.

Naturally, an FTA can be used to affect future policies only if its reversal is costly enough to inhibit future governments from reversing the arrangement. While here we simply assume that FTAs are irreversible, it would be relatively straightforward to extend the current model so that irreversibility becomes an equilibrium result, e.g. by relying on McLaren's (2002) notion that governments incur in "negotiating costs" when forming (or withdrawing from) FTAs. Ultimately, FTAs matter for commitment as long as there is a non-trivial cost to reverse them.<sup>15</sup>

It is also important to clarify that, although we consider a discrete-time problem, one should think of this as a continuous-time problem, where the realization of a coup depends on both exogenous (e.g. the state of the world economy) and endogenous (e.g. effort spent on coordination) factors. The point we develop below is that the latter is affected by policies such as the formation of an FTA. We keep the 2-period modeling only for expositional reasons.

### 3. THE DECISION TO FORM A FREE TRADE AGREEMENT

A free trade agreement between two countries is represented by the elimination of tariffs on each other's imports in all sectors included in the agreement. Thus, the equilibrium under an FTA is

---

<sup>15</sup> Irreversibility is also coherent with history, as preferential trading arrangements *de facto* implemented are seldom turned down later on. Even in the rather rare circumstances when authoritarian regimes gained control of a country that participated in an effective trade agreement, the arrangement is usually honored, as for example in Swaziland, a member of SACU.

analogous to the one described in Section 2, the only difference being the constraint imposed on some (potentially all) of the partners' reciprocal import tariffs. Without loss of generality, we let Home's potential FTA partner be country  $Y$ .

An FTA can be implemented by the incumbent government for reasons related or unrelated to the authoritarian threat. There are four possibilities. First, the country may already be a consolidated democracy, in the sense that condition (15) holds neither with nor without FTAs. This is the standard case considered in the regionalism literature, and it is not our goal to analyze it further here. Rather, we focus on situations where FTAs can be formed for "strategic" reasons.

The second possibility is that the country's democracy is so fragile that condition (15) is satisfied whether or not there is an FTA in place. In that case, while an FTA cannot be used to prevent a coup, the possibility of losing power can affect the incentives of the incumbent government with respect to the formation of the agreement. Finally, it is possible that an FTA affects (in either direction) the expected payoff of the authoritarian group and, as a result, its incentives to attempt to take power.

Before starting our analysis, we need however to describe the effects of an FTA on the level of available political rents and the role of parameter  $b$  in shaping its welfare effects. These results set the basis for the analysis of the political viability of FTAs.

### **3.A. The rent destruction effect**

Ornelas (2005a) shows that an FTA moderates the role of political economy forces in the determination of tariffs, and that the mitigation of the politically driven distortions corresponds to a source of welfare gain that is more relevant, the more far-reaching the government's political economy motivations. Furthermore, an FTA diminishes the rents created in the lobbying process. Intuitively, because the arrangement provides free access to the partner's exporters, the market share of the domestic industry shrinks, at any given external tariff. As a result, the FTA makes any price increase brought by a marginal increase in the external tariff less valuable for the import-competing industries, lowering their incentives to lobby for higher external tariffs. In equilibrium, these lower incentives result in a lower external tariff and in fewer rents for the government.<sup>16</sup> The following

---

<sup>16</sup> There is robust empirical evidence that the formation of free trade areas in developing countries (largely the focus of our analysis) leads to declining external tariffs (see Estevadeordal et al. 2008 for evidence on ten Latin America countries and Calvo-Pardo et al. 2011 for evidence on ten Southeast Asia countries), although the evidence is mixed for developed countries (see Limao 2006 and Bernhofen et al. 2012). While measuring protectionist rents directly is very difficult, the level of tariffs provides a good proxy for them; see Freund and Ornelas (2010) for a general discussion.

lemma summarizes these effects.<sup>17</sup>

**Lemma 1. The rent destruction effect of FTAs (Ornelas 2005a)**

Everything else constant, an FTA

- (a) improves Home’s welfare by more (or reduces it by less), the higher the government’s rent-seeking bias; and
- (b) reduces the political rents generated in the political process.

Lemma 1 allows us to analyze the conditions under which the Home government would choose to form an FTA.<sup>18</sup> The decision regarding the formation of an FTA is based on the anticipated impact of the agreement. The government implements the agreement if and only if it increases the government’s present value payoff. Note that the effects described in Lemma 1 are larger, the greater the number of specific-factor import-competing sectors included in the FTA.

Before proceeding to analyze how the possibility of political disruption affects the government’s willingness to form free trade agreements, let us define some useful notation. We henceforth attach subscript “*F*” to all variables when they are evaluated under an FTA. We adopt subscript “ $\Delta F$ ” to represent the *equilibrium change* in any variable due to the FTA. For example,  $W_{\Delta F}^x$  denotes the aggregate welfare change in a non-numeraire export sector due to the agreement, whereas  $W_{\Delta F}^m(b)$  and  $W_{\Delta F}^m(b=0)$  denote, respectively, the actual aggregate welfare impact of the FTA on a non-numeraire import sector and the equivalent effect under a hypothetical administration whose only concern is national welfare (equivalent to a situation where lobbying is effectively banned). Finally, let  $I^M$  denote the number of specific-factor import-competing sectors included in the FTA under analysis, with  $I^M \leq M$ . Aggregating the welfare impact of the agreement on both types of sectors, we then define  $W_{\Delta F}^M(b) \equiv I^M W_{\Delta F}^m(b)$  and  $W_{\Delta F}^X \equiv I^X W_{\Delta F}^x$ . Analogously,  $PR_{\Delta F} \equiv I^M PR_{\Delta F}^m$ .

**3.B. FTAs that do not affect the probability of political disruption**

We begin analyzing the case where there is a possibility of political disruption but this possibility is unaffected by the existence of FTAs – that is, condition (15) holds regardless of FTAs.

In this case, the equilibrium payoff of the incumbent democratic government under the FTA corresponds to

---

<sup>17</sup> These results do not hinge on the perfectly competitive structure adopted by Ornelas (2005a), which we follow here. Analogous results obtain also under oligopolistic competition (Ornelas 2005b).

<sup>18</sup> Naturally, an FTA is formed only if all prospective members endorse it. We conduct the discussion from the perspective of the Home country, but an analogous analysis would apply for country *Y*.



$$(16) \quad \Gamma_F^D = G_F + \delta[(1-\rho)G_F + \rho H_F].$$

The condition under which the democratic government supports the FTA when the authoritarian threat is inevitable is  $\Gamma_{\Delta F}^D \equiv \Gamma_F^D - \Gamma^D > 0$ . Using equations (12) and (16),  $\Gamma_{\Delta F}^D$  can be rewritten as

$$\Gamma_{\Delta F}^D = G_{\Delta F} + \delta[(1-\rho)G_{\Delta F} + \rho H_{\Delta F}].$$

Using expressions (9) and (10) and manipulating, this expression becomes

$$(17) \quad \Gamma_{\Delta F}^D = [1 + \delta(1-\rho)] \left[ W_{\Delta F}^M(b=0) + \frac{PR_{\Delta F}}{2} \right] + \delta\rho W_{\Delta F}^M(b) + (1+\delta)W_{\Delta F}^X.$$

Thus, the incumbent democratic government supports the FTA in this case if

$$(18) \quad [1 + \delta(1-\rho)][2W_{\Delta F}^M(b=0) + PR_{\Delta F}] + 2\delta\rho W_{\Delta F}^M(b) + 2(1+\delta)W_{\Delta F}^X > 0.$$

The interesting case is when the democratic government changes its stance toward an FTA *because* of the authoritarian threat. An FTA is (ordinarily) politically feasible if

$$(19) \quad 2[W_{\Delta F}^M(b=0) + W_{\Delta F}^X] + PR_{\Delta F} > 0.$$

The next proposition shows that the authoritarian threat can make an otherwise politically infeasible FTA (i.e., one for which condition (19) does not hold) into a politically feasible one.

**Proposition 1.** Even if the authoritarian threat cannot be affected, the mere possibility of political disruption can turn an otherwise politically unfeasible FTA into a viable one. By contrast, the possibility of disruption cannot render unfeasible an otherwise feasible FTA.

Proof: We need to show first that  $\Gamma_{\Delta F}^D$  increases with  $\rho$ . Using (19), we have that

$$(20) \quad \frac{d\Gamma_{\Delta F}^D}{d\rho} = \delta \left[ W_{\Delta F}^M(b) - W_{\Delta F}^M(b=0) - \frac{PR_{\Delta F}}{2} \right].$$

We know from Lemma 1 that  $PR_{\Delta F} < 0$  and that the welfare impact of an FTA is increasing in the rent-seeking bias of the government, so that  $W_{\Delta F}^M(b) - W_{\Delta F}^M(b=0) > 0$ . Accordingly, expression (20) is unambiguously positive, so  $\Gamma_{\Delta F}^D$  increases as the probability of disruption rises. As a result, an FTA that is politically unfeasible when there is no chance of political disruption can become viable if the likelihood of political disruption is high enough. That is, an FTA that does not satisfy condition (19) can satisfy criterion (18) for sufficiently high  $\rho$ . On the other hand, the reverse cannot happen: if an FTA is politically viable when there is no chance of political disruption, it remains viable if a possibility of change in power through force arises. That is, an FTA that satisfies condition (19) also satisfies criterion (18) for any  $\rho > 0$ . ■

Proposition 1 shows that the possibility of political disruption can enhance the political feasibility of FTAs by creating a “strategic” motivation for their adoption. Strategically supported FTAs arise when, between conditions (18) and (19), only the former is satisfied, so that

$$(21) \quad \Gamma_{\Delta F}^D(\rho = 0) \leq 0 < \Gamma_{\Delta F}^D(\rho > 0).$$

An FTA can be implemented for strategic reasons because the democratic government, if out of power, will not receive any of the lobbying-related rents. In that case, the government would benefit from an FTA because the agreement constrains the welfare-distorting political activities of the authoritarian group if it gets in power. Thus, a government that expects to lose power to a dictatorial group might seek an FTA simply to constrain the policies of the incoming authoritarian group. Since this strategic motivation is more relevant when disruption is more likely, it follows that “democratic instability” incites the formation of free trade agreements.<sup>19</sup>

The number of Home’s import-competing sectors susceptible to lobbying that are included in the FTA,  $I^M$ , also affects the possibility of strategically supported FTAs.

**Proposition 2.** The set of parameters under which the possibility of political disruption turns an otherwise unviable FTA into a politically viable one increases with the number of Home’s specific-factor import-competing sectors included in the agreement ( $I^M$ ).

Proof: To prove this result, it suffices to show that the probability of disruption,  $\rho$ , is a strategic complement of  $I^M$  in the function  $\Gamma_{\Delta F}^D$ , which gives the criterion for the political viability of FTAs. This function is represented in the left-hand side of (18). Based on the definition of welfare aggregated across all non-numeraire import sectors included in the agreement, we have that  $W_{\Delta F}^M(b) - W_{\Delta F}^M(b = 0) = I^M [W_{\Delta F}^m(b) - W_{\Delta F}^m(b = 0)]$ . Accordingly, the FTA affects political rents only in the included sectors:  $PR_{\Delta F} = I^M (PR_{\Delta F}^m)$ . Therefore, it follows that

$$\frac{d^2(\Gamma_{\Delta F}^D)}{d\rho dI^M} = 2\delta \left[ W_{\Delta F}^m(b) - W_{\Delta F}^m(b = 0) - \frac{PR_{\Delta F}^m}{2} \right] > 0.$$

Hence, the set of parameters under which condition (21) is satisfied enlarges as  $I^M$  increases. ■

Thus, the more comprehensive the FTA is, the greater is the extra incentive of the democratic

---

<sup>19</sup> It is easy to see that this strategic motivation for signing FTAs is stronger, the greater the rent-seeking bias of the autocrat. This follows because the forces underlined in Lemma 1 are stronger, the higher the rent-seeking bias of the group setting policies. This suggests that an authoritarian threat can make strategic FTAs particularly appealing, since despite some disagreement, the majority of views in the literature seem to agree that autocracies tend to pursue particularly distortionary policies (see Dixit 2010).

government to form the agreement because of the authoritarian threat. The reason is that, if Home imports more widely from its FTA partner in sectors where there is active lobbying, the agreement becomes more rent-destructing. While this is helpful for the country as a whole, it is detrimental to those in office who benefit from those rents. Under the threat of political disruption, however, the government understands that the loss of rents will be borne instead by the authoritarian group, if it is successful in gaining power. The destruction of rents is therefore less critical in the democratic government's evaluation of the agreement.

### 3.C. FTAs that can help secure democracies

The analysis above considers the case where a free trade agreement is not pivotal in the decision of the authoritarian group to attempt to take power through force. But this need not be the case. We now show that an FTA can change the sign of condition (15). However, the change can go in only one direction. Specifically, an FTA can prevent a coup from happening, but it cannot provoke a coup that would not occur without the agreement.

**Proposition 3.** If the authoritarian group did not intend to initiate a coup in the absence of trade agreements, an FTA cannot induce it to initiate one. On the other hand, the formation of a sufficiently rent-destructing FTA can free the country from the authoritarian threat. This is more likely to happen, the greater the number of Home's specific-factor import-competing sectors included in the agreement ( $I^M$ ).

Proof: In the absence of trade agreements, the authoritarian group attempts to take power through a coup if condition (15) is satisfied. With an FTA, a similar condition applies:

$$(22) \quad [W_F^M(b=0) - W_F^M(b)] + \frac{PR_F}{2} > \frac{1 - \rho(DC)}{\rho(DC)} K.$$

The difference between conditions (15) and (22) is in the expressions' left-hand sides, which denote the gains of the authoritarian group from getting power. On the other hand, the FTA impacts neither the probability of success of a coup nor the costs of a failed coup attempt. Subtracting the left-hand side of inequality (15) from the left-hand side of inequality (22), we obtain

$$(23) \quad [W_{\Delta F}^M(b=0) - W_{\Delta F}^M(b)] + \frac{PR_{\Delta F}}{2} < 0,$$

where the negative sign follows directly from Lemma 1. Thus, if condition (15) is not satisfied, condition (21) will not be satisfied either, implying that an FTA cannot provoke a coup that otherwise would not occur. Conversely, condition (15) can be satisfied while condition (21) is not,

implying that an FTA can be critical to prevent the authoritarian group from seeking power. Finally, note that the left-hand side of (23) decreases with  $I^M$ :

$$\frac{d[\text{lhs}(25)]}{dI^M} = [W_{\Delta F}^m(b=0) - W_{\Delta F}^m(b)] + \frac{PR_{\Delta F}^m}{2} < 0.$$

Thus, the range of parameters under which an FTA is pivotal in preventing the authoritarian threat is larger, the greater the number of non-numeraire import-competing sectors the FTA includes. ■

Proposition 3 shows that, because of the rent-destructing effects of FTAs, a free trade agreement can critically reduce the incentives of the authoritarian group to attempt to subvert the country's democratic system. In this sense, an FTA can help to constrain the emergence of authoritarian regimes, especially if the bloc is significantly rent-destructing, as in that case it will be more effective in lowering the gains from power of the authoritarian group. Relying on the common notion that the availability of rents can entice political turbulence—while the unavailability of rents can prevent it—the proposition's novelty stems from the recognition of free trade agreements as instruments to restrain the gains from rent-seeking behavior.

We still need to ask, however, whether the incumbent democratic government would actually *want* to implement the arrangement. The next proposition shows that the possibility of using an FTA to block a coup necessarily raises the government's incentives to sign an agreement.

**Proposition 4.** An FTA can become politically feasible by being pivotal to prevent a coup.

Proof: When an FTA cannot prevent the authoritarian group from seeking power through force, it is politically viable if  $\Gamma_{\Delta F}^D = G_{\Delta F} + \delta[(1-\rho)G_{\Delta F} + \rho H_{\Delta F}] > 0$ . When the agreement reverses the decision of the authoritarian group, it is adopted by the democratic government if

$$G_F + \delta G_F > G + \delta[(1-\rho)G + \rho H],$$

where the left-hand side represents the present value of the government under the agreement (and no authoritarian threat) and the right-hand side corresponds to its expected present value without the FTA (and with the authoritarian threat). This condition can be rewritten as

$$(24) \quad (1 + \delta)G_{\Delta F} + \delta\rho(G - H) > 0.$$

Now notice that the left-hand side of (24) is greater than  $\Gamma_{\Delta F}^D$  if  $G_F > H_F$ , which is true from the definitions of  $G_F$  and  $H_F$ , which are analogous to those in (9) and (10). Hence, even if  $\Gamma_{\Delta F}^D < 0$ , condition (24) can be satisfied. ■

Proposition 1 asserts that the possibility of political disruption can render feasible an

otherwise unfeasible free trade agreement. Proposition 4 indicates that the political support for an FTA is further enhanced if the agreement can also play a role in preventing disruption of the political system. This is true even though here we abstract from any ideological motivation the incumbent democratic government may have; if the government perceived a benefit *per se* from maintaining democracy in the country, its incentive to form an FTA that can serve that purpose would be further enhanced.

It is also worth noting that, although we abstract from direct lobbying for and against FTAs, adding *ex-ante* lobbying would have little effect on the analysis, just as Ornelas (2005a) shows to be the case when changes in the political regime are not considered. Consider for example exporting sectors, the main source of support for FTAs. Their benefits from an FTA depend primarily on the extent of the access to the partners' markets, rather than on domestic policies. Therefore, their willingness to support an FTA is not directly affected by who is in power.

Our results thus suggest that free trade agreements—especially those that are particularly effective in destroying rents—can be useful to prevent an authoritarian threat. This can be especially important in fledgling democracies, given the instability that typically follows the end of dictatorial regimes. We now turn to showing that these relationships are also empirically robust.

#### 4. EMPIRICAL STRATEGY

The model has two main predictions about the relationship between FTAs and democracy, which imply the following hypotheses:

- H1. Participation in FTAs lowers the probability of democratic failure.
- H2. Unstable democracies are more likely to form FTAs.

To test H1, our dependent variable is a dummy indicating whether democracy was interrupted, or alternatively the length of democratic spells. This allows us to estimate the probability that democracy will fail in the country, which we denote by  $\text{Prob}(\text{enddemo})$ . We define democracy failure in different ways, based alternatively on Polity IV data and on a dichotomous classification from Cheibub et al. (2010); we explain this in detail in Section 5. The key independent variable is a measure of the intensity of the country's participation in FTAs.

To test H2, our dependent variable is the change in a country's FTA participation. The key independent variable is a measure of democratic instability that reflects the *expectation* that the democratic regime may fail in the country.

As indicated in the Introduction, our problem is related to the one studied by Persson and Tabellini (2009), who examine the determinants (in particular the effect of income) of the stability of democracies and the impact of this perceived stability on income growth. Our empirical strategy resembles their approach.

#### 4.A. Testing H1: Participation in FTAs and democracy survival

We estimate the likelihood of democratic failure relying on the concepts of domestic democratic capital (DOM) and foreign democratic capital (FOR) developed by Persson and Tabellini (2009), while adding a variable that captures the intensity of a country's participation in FTAs. DOM is a measure of the democratic history of the country, whereas FOR measures current levels of democracy in the world.<sup>20</sup> Other explanatory variables include economic factors (e.g. GDP per capita, denoted by vector X) and geographical and institutional factors (e.g. war indicators, continent of location and legal origin, denoted by vector Z).

The dataset covers only countries' democratic spells. We estimate a discrete time duration analysis modeled as logit, which can be implemented as follows:

$$(25) \quad \log(P/(1-P)) = \alpha_0 + \alpha_1 \text{FTA}_{t-1} + \alpha_2 \text{DOM}_{t-1} + \alpha_3 \text{FOR}_{t-1} + \alpha_4 X_{t-1} + \alpha_5 Z_{t-1} + u_t$$

where P denotes Prob(enddemo) and subscript t-1 indicates that the variable is lagged one year.<sup>21</sup>

The variable FTA in equation (25) represents the extent (or the "intensity") of a country's participation in FTAs in a given year. Measuring the FTA intensity of a country is far from trivial. Despite a prolific literature on the consequences of preferential trade integration, that line of research offers no guidance on how to measure this intensity. In fact, most empirical regionalism papers simply use dummies to represent FTA participation. While this may be adequate for other purposes, such a measure is inappropriate here for several reasons. First, unlike other studies where the unit of observation is a country dyad, we need a measure of FTA participation at the *country* level, since we want to estimate the endurance of democracy in individual countries. And while some countries participate in a single (or no) FTA, others are members of multiple agreements. Second, there is wide heterogeneity among FTAs. While some arrangements are fully implemented, others are not, implying that preferences actually offered are few and small, and therefore entail little destruction of rents. Furthermore, some agreements are very large (e.g. the European Union),

---

<sup>20</sup> In the next section we provide a precise definition of both variables.

<sup>21</sup> An alternative to this logit specification would be a complementary loglog (cloglog) regression, where we treat the time interval as discrete or grouped by year. However, when the probability of positive outcomes is small as in our case, where democratic failures account for only 2.2% of the sample (see Table 1), a cloglog link function is similar to a logistic link function. Since logistic regressions are more conventional than cloglog ones, we focus on the former. Results from cloglog regressions are similar.

while others are tiny, including some that have many members (e.g. CARICOM). Third, even within a given FTA, the impact of the bloc can be very different on each member. Consider NAFTA: while it has a very large impact on Mexico, the smallest member, its effects are much less pronounced in the United States, the largest member. Now, according to Proposition 3, only sufficiently rent-destructing FTAs have an effect on the sustainability of democracy. We clearly need, therefore, a more precise measure of the intensity of a country's participation in FTAs than what FTA dummies can offer.

In our model, where all sectors are symmetric, the extent of rent destruction within an FTA is given by the number of import-competing sectors included in the agreement. More generally, it depends also on the size of the export sectors of the FTA members relative to Home's import-competing sectors. To capture both, we use in our main regressions the share of imports from FTA members. This variable does not indicate whether the imports from the FTA partners are in sectors where lobbying takes place, which is where the rent destruction effect of FTAs will be operative. In fact, identifying empirically those sectors in every country would be virtually impossible. Nevertheless, the variable has the central virtue of varying monotonically with the degree of implementation of the agreement and with the importance of the agreement for the country in question. Consequently, it should be positively correlated also with the variable  $I^M$  in our model, which represents the extent to which import-competing sectors where lobbying happens are included in the agreement. Hence, the import share from FTA members provides a useful proxy for the country-level degree of rent destruction engendered by the FTAs a country belongs to,<sup>22</sup> and it appears to us to be much superior to alternative methods of measuring FTA participation available in the literature (almost always dummy variables).

We also need to take into account the possibility of duration dependence in (25), i.e. the extent to which the conditional hazard of democracy rises or falls over time. If there is duration dependence, the hazard of enddemo will depend on the duration of the democratic regime. In general, its effect can be either positive or negative. Domestic democratic capital will capture part of the duration dependence. For the residual duration dependence, we use a polynomial of a time counter that counts the number of years passed since the beginning of the current democratic spell. The order of the polynomial is determined by the best fit in the regressions.<sup>23</sup> Including this duration polynomial, which we denote by DUR, we rewrite our estimating equation as

---

<sup>22</sup> We use a country's import/GDP ratio to control for the overall importance of its import sector.

<sup>23</sup> Using year dummies to account for duration dependence is not an appealing alternative for our analysis, since all the years without a democratic failure would be dropped from the sample, causing the loss of important cross-sectional variation.

$$(26) \quad \log(P/(1-P)) = \alpha_0 + \alpha_1\text{FTA}_{t-1} + \alpha_2\text{DOM}_{t-1} + \alpha_3\text{FOR}_{t-1} + \alpha_4X_{t-1} + \alpha_5Z + \alpha_6\text{DUR} + u_t.$$

A remaining concern in (26) is unobserved heterogeneity. It is possible that some countries are more likely to have interrupted democracies due to unobserved variables that are correlated with some right-hand side variables in (26). To deal with this possibility, we run a country random effects logit specification.<sup>24</sup>

Alternatively, we also use continuous time duration analysis for equation (26), defining the dependent variable as the duration of a democracy spell, i.e. the number of years passed since the onset of each democracy spell until it was interrupted or right-censored.<sup>25</sup> This variable is the same as the time counter we use for the duration dependence in the logit specifications. Hence the unit of analysis here is a democratic spell. The duration dependence is specified parametrically (Weibull model) or non-parametrically (Cox proportional hazard model). In the Weibull model, the hazard function is  $h(t) = \varphi \exp(\beta'X_t)t^{\varphi-1}$ , where  $\varphi$  is a shape parameter to be estimated and  $t$  is the duration time. In the Cox model, the hazard is  $h(t) = h(0)\exp(\beta'X_t)$ , where the baseline hazard,  $h(0)$ , is allowed to be group-specific.

#### 4.B. Testing H2: Democracy instability and FTA formation

Once we have the predicted  $\text{Prob}(\text{enddemo})$ , we can use it to test our second hypothesis, that the likelihood of democratic failure helps to explain the formation of FTAs. In the analysis we also include the economic, geographical and institutional variables used in the duration regression as controls, except DOM, FOR and the duration dependence terms. Nothing in our theory suggests that these variables should have an independent effect on the change of a country's FTA participation,  $\Delta\text{FTA}$ , in addition to their indirect effects on  $\Delta\text{FTA}$  through  $\text{Prob}(\text{enddemo})$ . This is analogous to the identification assumption of Persson and Tabellini (2009), that democratic capital affects income growth through their effects on the sustainability of democracy only. We also exclude the duration dependence terms; instead, we use year dummies to capture time effects. In our context, the rationale to exclude FOR may not apply, because the level of democracy in a region could have an independent effect on the likelihood that countries in the region will form FTAs. Thus, for

---

<sup>24</sup> Acemoglu et al. (2008) highlight the importance of including country fixed effects when studying the effect of income on democracy. Unlike in that type of analysis, in our case a fixed effects logit procedure (i.e. conditional logit) would be inappropriate because most countries do not experience democracy failure during our sample period. As a result, the observations for all long-lived regimes would be dropped, eliminating much of the cross-sectional variation in the data that helps us to capture the effect of FTAs on democracy survival.

<sup>25</sup> Since the Polity dataset can go as far back as 1800, we can determine the onset of the democratic spell even for most countries that have a democratic regime in the beginning of our sample (1960). In the few instances when this is not possible, we set the onset of the democratic spell to the first year for which the polity score is available.



robustness we also run a specification where we include in the  $\Delta$ FTA regression both DOM and FOR, as well the duration dependence terms. In that case we rely only on the non-linearity of our regression to identify the parameters of interest. In all cases, in line with our model we use import shares from FTA partners to measure the intensity of a country's FTA participation. We also include country fixed effects in all regressions.<sup>26</sup>

To test H2, we then run the following specification:

$$(27) \quad \Delta\text{FTA} = \beta_0 + \beta_1\text{Prob}(\text{enddemo})_{t-1} + \beta_2\text{FTA}_{t-1} + \beta_3X_{t-1} + \beta_4Z_{t-1} + \{\text{country and year FEs}\} + v_t.$$

In equation (27) we measure  $\text{Prob}(\text{enddemo})$  with its predicted hazard rate evaluated at  $\text{FTA}_{t-1}=0$ , which we denote by  $\hat{h}$ . We do so to eliminate the effect of FTAs on our measure of political instability (we already include  $\text{FTA}_{t-1}$  itself in (27)). We also include the squared term of  $\hat{h}$  to capture possible nonlinearities. Since  $\hat{h}$  is a constructed regressor, we adjust standard errors using bootstrapping methods. Our model predicts a positive impact of  $\hat{h}$  on FTA formation.

## 5. DATA

We have a panel with 116 countries that have experienced democracy at some point during 1960-2007. Not every country is included because our study is restricted to democracies. In the Online Appendix ([http://personal.lse.ac.uk/ornelas/Liu&Ornelas\\_Appendices.pdf](http://personal.lse.ac.uk/ornelas/Liu&Ornelas_Appendices.pdf)) we list the countries covered in our enddemo duration analysis. More than 200 countries are covered in the construction of our FTA measures, as explained below.

An unavoidable difficulty in any empirical analysis of regional trade agreements is that some of them represent little more than scraps of paper. We try to turn this problem to our favor, by examining whether our predictions hold similarly for the shallowest arrangements; if the mechanism at work is indeed the destruction of rents, it should not.

One still needs, however, to determine which agreements are those. To make this distinction, we follow a simple set of rules. First, we classify as *free trade agreements* all free trade areas and customs unions ratified under GATT's Article XXIV, which we refer to as full-fledged free trade agreements (or FTAs, for brevity). The Article specifies minimum liberalization requirements that preferential trade agreements have to satisfy. Developed countries must notify their agreements to

---

<sup>26</sup> Even though much has been written about regionalism, we still know relatively little about what makes governments willing to form FTAs. Consider for example the seminal contribution by Baier and Bergstrand (2004). Most of the explanatory variables they consider are either geographical, which do not change over time, or "structural," in the sense of changing little over time (e.g. factor endowments). Since we work with a panel, our fixed effects capture all of those fixed/almost fixed factors.

the WTO under Article XXIV. Trading blocs formed only by developing economies can notify either under Article XXIV or under the Enabling Clause, which imposes much fewer constraints on what bloc members must accomplish. Accordingly, we consider agreements notified under the Enabling Clause—as well as those not notified to the WTO—to be *partial-scope preferential trade agreements* (PTAs, in short). These rules, while sensible, tend to overstate the number of PTAs relative to FTAs, because some developing countries may choose to notify the agreement under the Enabling Clause (to gain flexibility in their liberalization path) even when they intend to liberalize significantly vis-à-vis each other. To avoid this bias, we delved into the literature studying specific agreements to identify such cases. We identified three agreements—Mercosur, CAN and ASEAN—that are notified under the Enabling Clause but have implementation rates comparable to several Article XXIV FTAs. We could not find studies providing similar evidence for other Enabling Clause PTAs. Importantly, eliminating those exceptions does not have a qualitative important effect on our estimates.

Data for the agreements come from the WTO website and also from information available in several other sources. In the Online Appendix, we list the agreements in our dataset with their types (FTAs or PTAs) and other information about the data sources. As discussed in the previous section, we proxy a country's intensity of FTA participation by its imports from FTA partners as a share of its total imports in a given year. We use an analogous definition for PTA participation:

- FTA\_impsh: a country's imports from FTA partners as a share of its total imports;
- PTA\_impsh: a country's imports from PTA partners as a share of its total imports.

As shown in Table 1, the average import share from PTAs in our sample is 0.06, compared to 0.21 for FTAs. We obtain the import data from the IMF Direction of Trade Statistics. To construct the shares, we carefully consider the dates of the formation of new blocs, of the accession of new members, and of the de-activation of existing blocs.

As we follow Persson and Tabellini's (2009) general estimation strategy, our first definition of democracy failure follows their definition, which relies on Polity IV data (available at <http://www.systemicpeace.org/polity/polity4.htm>). This entails defining a regime as "democratic" iff its polity2 score (which ranges from -10 to 10, with higher values representing more democratic regimes) is strictly positive.<sup>27</sup> In a democratic spell, enddemo is zero as long as democracy remains uninterrupted and becomes unity when it ends. If a democracy does not end during our sample, enddemo is right-censored. There are 61 episodes of enddemo in our sample according to this definition (see the Online Appendix for the list of episodes). For those transitions, the median score

---

<sup>27</sup> Countries enter the sample as they become independent, but only if they have a strictly positive polity2 score, since our study is restricted to democracies.

before the change is 4.8, whereas the median score after the change is -3.5; the median drop in the polity2 score is 8.8 points. A representative example of the median case of democracy failure is Thailand from 1970 to 1971, when its polity score fell from 2 to -7.

Table 1: Descriptive statistics of main variables

Variable	Description	Mean	S.D.	Min	Max
enddemo	Dummy indicating the end of a democracy	0.022	0.145	0	1
L.FTA_impsh	Lagged import share from FTA partners	0.214	0.282	0	0.920
L.PTA_impsh	Lagged import share from PTA partners	0.063	0.128	0	0.698
$\Delta$ FTA_impsh	Change in FTA_impsh from previous year	0.010	0.067	-0.451	0.761
$\Delta$ PTA_impsh	Change in PTA_impsh from previous year	0.002	0.032	-0.442	0.502
L. log(GDP/capita)	Lagged log(GDP/capita)	8.004	1.522	4.400	10.632
L.DOM	Lagged current domestic democratic capital	0.573	0.344	0	1.000
L.FOR	Lagged foreign democratic capital	0.041	0.140	-0.217	0.261
war	Current war indicator	0.063	0.243	0	1
L.war	Lagged war indicator	0.066	0.249	0	1
socialist	leg_origin = Socialism	0.109	0.312	0	1
africa	Africa dummy	0.169	0.375	0	1
middleeast	Middle East region dummy	0.032	0.176	0	1
Spain_colony	UK colony dummy	0.189	0.391	0	1
UK_colony	Spain colony dummy	0.307	0.461	0	1
L.(M/GDP)	Lagged import/GDP	0.358	0.197	0.036	1.528
duration	Duration of democracy (# of years passed)	37.955	43.882	1	208
Hazard	Predicted hazard	0.025	0.038	0	0.386
enddemo06	enddemo with stricter thresholds	0.011	0.102	0	1
enddemo64	enddemo with stricter thresholds	0.018	0.132	0	1
L.FTA_impsh_moredemo	FTA_impsh with more democratic partners	0.056	0.165	0	0.899

Note: The descriptive statistics of most of the variables in this table are based on the sample with 2,827 observations used in regressions (2) to (5) in Table 2A, except for the last three variables which are based on regressions (1) and (3) in Table 2C and the second regression in Table 4, respectively.

We also use a different measure of democracy failure from a recent database developed by Cheibub et al. (2010, available at [https://netfiles.uiuc.edu/cheibub/www/DD\\_page.html](https://netfiles.uiuc.edu/cheibub/www/DD_page.html)).<sup>28</sup> Their measure has the advantage of offering a dichotomous classification of democracy/autocracy that yields a straightforward definition of the transitions—unlike the Polity index, where transition needs to be defined according to (necessarily) arbitrary thresholds. There are 44 cases of enddemo in our sample according to this definition (see Online Appendix).

Following Persson and Tabellini (2009), the construction of DOM and FOR is also based on polity2 scores. DOM is defined as

$$DOM_{it} = (1 - \delta) \sum_{\tau=0}^{\tau=t-t_0} d_{i,t-\tau} \delta^\tau,$$

<sup>28</sup> Cheibub et al. (2010) extend the classification proposed initially by Alvarez et al. (1996).

where  $\delta$  is a discount factor and  $d_{i,t-\tau}$  is a dummy for a strictly positive polity2 score. The first positive dummy for each country is either the first year with a strictly positive polity2 score or the first year the country appears in the Polity dataset, which for some countries goes back to 1800. As Persson and Tabellini, we find that what really matters for democratic stability in DOM is *current* DOM (i.e. the current democratic spell), whereas the portion of DOM due to previous democratic spells is usually insignificant in the regressions. Accordingly, we use current DOM in all of our regressions, so  $t_0$  corresponds to the first year in which  $d_{it} = 1$  in the current democratic spell. For the discount factor, we adopt  $\delta = .95$ ; results change little for  $\delta \in [.94, .99]$ , the range considered by Person and Tabellini. In turn, FOR is defined as

$$FOR_{it} = \sum_{j \neq i} Polity_{jt} \left( 1 - \frac{Dist_{ij}}{DEq} \right) / N_t ,$$

where  $Polity_{jt}$  is country  $j$ 's polity2 score at  $t$  (rescaled to the [-1,1] interval),  $Dist_{ij}$  is the distance between the capitals of countries  $i$  and  $j$ ,  $DEq$  is half the length of the equator, and  $N_t$  is the number of independent countries in the world at  $t$ . Thus, the closer a country is to other democracies, the greater its own stock of foreign capital.

GDP per capita data come from the World Development Indicators database. Data on wars come from the Correlates of War dataset and includes all wars a country was involved in. Legal origin data are drawn from La Porta et al. (1999). Colonial history variables come from the CIA's World Fact Book. WTO membership data come from the WTO website. Trade openness measures are obtained from the Penn World Table 6.3. The data used to calculate the number of international organizations (IO) come from the database for International Governmental Organizations (IGO, v2.3).<sup>29</sup> The data on formal military leader as chief executive are based on Gandhi and Przeworski (2006).<sup>30</sup> Table 1 lists the definitions of all the variables used in the main regressions and provides descriptive statistics.

## 6. EMPIRICAL RESULTS

### 6.A. Does participation in FTAs affect democracy survival?

We study first the impact of lagged FTA participation on the duration of democracy. Table 2A shows the duration analysis results for five different specifications. The logit regression (1) uses only the FTA import share variable as a regressor. This variable alone explains around 7% of the variation

---

<sup>29</sup> Available at <http://www.correlatesofwar.org/COW2%20Data/IGOs/IGOV2.3.htm>.

<sup>30</sup> We thank James Raymond Vreeland for kindly sharing the data with us.

in enddemo, as shown by the Pseudo R2. In regression (2) we add all control variables and the duration dependence terms. It turns out that a second-order polynomial of the time counter produces the best fit of the model. In regression (3) we use country random effects, yet the LR test of the random effects (“rho”) is insignificant at the 10% level and the changes in the estimated coefficients are overall quite small. Standard errors in those regressions are clustered at the country level. Columns (4) and (5) show results for the continuous time duration models, where the dependent variable measures the number of years passed since the onset of a democratic regime. Column (4) shows the coefficient estimates of the Weibull regression, whereas column (5) displays the coefficients of the PH Cox model.

In all five specifications, the coefficient of the FTA import share is negative and statistically significant at the 5% level. This result supports our first hypothesis that greater participation in FTAs lowers the probability of democratic failure in a country. As for the control variables, GDP/capita, DOM, FOR and UK colony dummy have a negative and statistically significant impact on enddemo, while the war dummy has the opposite effect. The estimated coefficients of the other variables are statistically insignificant.

Table 2B reports estimates when we use the Cheibub et al.’s (2010) definition of democratic failure. Duration dependence terms are updated accordingly. Results are qualitatively very similar to those from Table 2A, although the estimated coefficients are larger. This reinforces the evidence that more FTA participation tends to increase the longevity of democracies.

Are the estimates also economically significant? It is useful to interpret the coefficients in terms of marginal effects. Consider the logit regression result in column (2) of Table 2B. The corresponding average marginal effect of FTA\_impsh is 0.06. It implies that a one standard deviation increase in the FTA import share (0.282) would decrease the hazard of enddemo by about 1.7 percentage points ( $= 0.282 \times 0.06 \times 100$ ) on average, which is about 67% of the average predicted hazard (2.5 p.p.). Compared with the variables Persson and Tabellini (2009) put forward as key determinants of democracy survival, this reduction is greater (in absolute terms) than that following a one-standard-deviation increase in foreign democratic capital (1.3 p.p.), though smaller than the corresponding figure for domestic democratic capital (8 p.p.). Or consider Mongolia, the only country without any FTA by 2010. According to the logit estimate, Mongolia’s hazard rate in 2005 would drop from 4.76% to 2.42% if it had the same FTA import share as Chile (FTA\_impsh=0.39 in 2005), or to just 0.4%, if Mongolia had the same FTA import share as Mexico (FTA\_impsh=0.72 in 2005).

Table 2A: Enddemo regression results using Polity data, FTAs

	(1)	(2)	(3)	(4)	(5)
	logit	logit	xtlogit	Weibull	PH Cox
<b>L.FTA_impsh</b>	<b>-5.907***</b> <b>(1.864)</b>	<b>-3.027**</b> <b>(1.335)</b>	<b>-3.131**</b> <b>(1.427)</b>	<b>-3.410**</b> <b>(1.381)</b>	<b>-2.938**</b> <b>(1.317)</b>
L.log(GDP/capita)		-0.558*** (0.150)	-0.628*** (0.239)	-0.741*** (0.153)	-0.569*** (0.149)
L.DOM		-11.935** (5.583)	-11.783 (8.839)	-17.830*** (4.642)	-71.748 (0.000)
L.FOR		-4.723*** (1.569)	-5.116*** (1.730)	-5.951*** (1.607)	-5.028*** (1.535)
war		1.285* (0.708)	1.297** (0.643)	1.258* (0.672)	0.972* (0.545)
L.war		-0.290 (0.689)	-0.323 (0.681)	-0.677 (0.701)	-0.292 (0.602)
Socialist		0.256 (0.587)	0.395 (0.707)	0.270 (0.621)	0.201 (0.583)
africa		0.395 (0.406)	0.456 (0.466)	0.050 (0.359)	0.270 (0.361)
middleeast		0.255 (1.190)	0.347 (1.179)	0.978 (0.895)	0.444 (1.049)
Spain_colony		-0.485 (0.486)	-0.470 (0.490)	-0.450 (0.471)	-0.451 (0.449)
UK_colony		-1.216** (0.477)	-1.301** (0.507)	-1.233*** (0.426)	-1.229*** (0.406)
L.(M/GDP)		-0.032 (1.042)	-0.032 (0.890)	0.083 (0.981)	0.095 (0.914)
Duration		0.613** (0.260)	0.631 (0.411)		
Duration^2		-0.009** (0.004)	-0.010* (0.006)		
Country RE			Yes		
rho			0.060		
Test rho=0 (p-value)			[0.323]		
Pseudo R2	0.070	0.211			0.111
Log Lik	-309.2	-232.2	-232.1	-97.15	-235.4
Observations	3,008	2,827	2,827	2,827	2,827

Notes: Robust standard errors in parentheses, clustered by country in logit regressions.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

In Table 2C we show that, when using the polity2 scores to define transition, the precise level of the threshold is not critical to our qualitative results. Surely, when defining democracy failure by a move in the polity2 score from strictly positive to non-positive, we may capture some very minor political changes (say a change in the score from 1 to 0). Of course we also use the binary measure from Cheibub et al. (2010), which is not open to this type of criticism, but it is nevertheless useful to check whether such marginal changes in democratic status are driving our results with the Polity data.

Table 2B: Enddemo regression results using Cheibub et al.'s (2010) democracy data, FTAs

	(1)	(2)	(3)	(4)	(5)
	logit	logit	xtlogit	Weibull	PH Cox
<b>L.FTA_impsh</b>	<b>-8.144***</b> <b>(2.265)</b>	<b>-4.963**</b> <b>(1.930)</b>	<b>-5.262*</b> <b>(3.164)</b>	<b>-4.912***</b> <b>(1.804)</b>	<b>-3.986**</b> <b>(1.890)</b>
L.log(GDP/capita)		-0.400* (0.209)	-0.441* (0.229)	-0.666*** (0.223)	-0.548*** (0.196)
L.DOM		-2.040 (1.292)	-2.197 (1.506)	-3.374*** (1.137)	-1.437 (1.352)
L.FOR		-5.566*** (1.688)	-5.800*** (1.908)	-7.117*** (2.008)	-6.184*** (1.996)
war		0.772 (1.062)	0.744 (0.816)	1.069 (0.669)	1.387** (0.646)
L.war		-0.652 (1.041)	-0.622 (0.872)	-1.138* (0.649)	-1.426** (0.659)
Socialist			-18.161 (8,057)	-15.061*** (0.875)	-44.731 (0.000)
africa		1.334** (0.654)	1.414** (0.633)	1.555** (0.687)	1.053* (0.567)
middleeast			-18.281 (16,401)	-13.799*** (1.341)	-43.373 (0.000)
Spain_colony		0.741 (0.531)	0.764 (0.492)	0.919 (0.596)	0.393 (0.534)
UK_colony		-0.833 (0.517)	-0.938 (0.585)	-1.134** (0.557)	-0.965* (0.539)
L.(M/GDP)		-0.557 (1.491)	-0.572 (1.279)	0.418 (1.136)	0.053 (1.126)
Duration		0.132** (0.062)	0.157* (0.092)		
Duration^2		-0.003** (0.001)	-0.003* (0.002)		
Country RE			Yes		
rho			0.073		
Test rho=0 (p-value)			[0.255]		
Pseudo R2	0.085	0.211			0.193
Log Lik	-248.1	-168.9	-168.8	-63.48	-116.0
Observations	3,279	2,122	2,484	2,396	2,396

Notes: Robust standard errors in parentheses, clustered by country in logit regressions.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

In columns (1)-(3) of Table 2C, we no longer classify as democracy failure the cases where the polity2 score drops by less than 6 points in the 3 years following the last year with a strictly positive polity score.<sup>31</sup> In columns (4)-(6), we follow Mansfield and Pevehouse (2006) and set the threshold for democratic status at a higher level, 7, and again we classify as democracy failure only

<sup>31</sup> Because the country random effects component in column (3) of Tables 2A and 2B are not significant, as shown by the estimated "rho" and its p-values, we use logit without random effects in Table 2C and some of the other robustness checks reported later.

cases where there is a non-trivial drop (of at least 4 points) in the polity2 score. With either of these stricter definitions of democracy, the coefficient estimates remain statistically significant; they also nearly double in size.

Table 2C: Robustness checks, enddemo regressions with **stricter thresholds**

	Enddemo06			Enddemo64		
	(1)	(2)	(3)	(4)	(5)	(6)
	logit	Weibull	PH Cox	logit	Weibull	PH Cox
<b>L.FTA_impsh</b>	<b>-6.261**</b> <b>(3.007)</b>	<b>-6.332**</b> <b>(3.102)</b>	<b>-6.188**</b> <b>(2.846)</b>	<b>-5.408**</b> <b>(2.176)</b>	<b>-5.897***</b> <b>(2.155)</b>	<b>-5.779***</b> <b>(2.019)</b>
L.log(GDP/capita)	-0.624** (0.255)	-0.796*** (0.204)	-0.708*** (0.234)	-0.610** (0.247)	-0.923*** (0.296)	-0.708*** (0.266)
L.DOM	-3.133 (4.177)	-11.469** (4.599)	-6.391* (3.591)	-5.618 (6.142)	-18.986* (11.031)	-4.063 (5.836)
L.FOR	-2.394 (3.001)	-4.050 (2.875)	-2.861 (2.666)	-1.181 (3.116)	-2.661 (3.147)	-1.790 (3.069)
war	1.419* (0.747)	1.304* (0.788)	1.831** (0.833)	0.954 (0.851)	0.780 (0.972)	0.340 (1.279)
L.war	-0.809 (1.301)	-0.822 (1.018)	-0.864 (0.910)	0.948 (0.948)	0.662 (0.841)	1.955 (1.345)
socialist	-1.140 (1.217)	-1.029 (1.341)	-1.238 (1.199)	-0.983 (0.939)	-0.735 (0.943)	-1.047 (0.989)
africa	-0.280 (0.739)	-0.554 (0.727)	-0.629 (0.691)	-0.536 (0.652)	-0.703 (0.597)	-0.748 (0.611)
middleeast		-13.573*** (1.072)	-43.489 (0.000)		-14.263*** (1.108)	-44.358 (0.000)
Spain_colony	-1.985** (0.878)	-2.092** (0.886)	-2.177** (0.870)	-2.861*** (0.867)	-2.628*** (0.999)	-3.266*** (1.050)
UK_colony	-1.068 (0.871)	-1.335 (0.853)	-1.188 (0.798)	-1.100 (0.759)	-1.488* (0.777)	-1.223 (0.808)
L.(M/GDP)	0.812 (1.703)	0.125 (2.049)	0.858 (1.930)	0.983 (1.434)	1.575 (1.311)	1.290 (1.357)
Duration	0.133 (0.201)			0.389 (0.258)		
Duration^2	-0.003 (0.003)			-0.009* (0.005)		
Pseudo R2	0.259		0.237	0.288		0.254
Log Lik	-69.00	-36.56	-49.22	-82.90	-33.90	-58.44
Observations	1,986	2,063	2,063	1,986	2,060	2,060

Notes: Regressions (1)-(3) consider only the cases of enddemo where the polity2 score transits from strictly positive to non-positive and drops by at least 6 points within a 3-year window since the transition. Regressions (4)-(6) consider only the cases of enddemo where the polity2 score transits from 7 or higher to 6 or lower and drops by at least 4 points within a 3-year window since the transition. Robust standard errors in parentheses, clustered by country in logit regressions. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3 shows in turn the results for partial-scope PTAs, using both Polity and Cheibub et al.'s definitions of democratic transition. If the mechanism through which trade agreements help democracies to consolidate is the destruction of protectionist rents, as proxied by countries' FTA



shares, then we should detect an effect of FTAs on democracy survival but not of PTAs. Indeed, when adding the PTA import share to the regression, its coefficient is always statistically indistinguishable from zero (while the coefficient on FTA\_impsh hardly changes), confirming that partial, incomplete processes of preferential trade liberalization have no meaningful effect on democracy survival.<sup>32</sup>

Table 3: Enddemo regression results, **partial-scope PTAs**

	Enddemo based on Polity			"tta" based on Cheibub et al.		
	(1)	(2)	(3)	(4)	(5)	(6)
	logit	Weibull	PH Cox	logit	Weibull	PH Cox
L.FTA_impsh	<b>-2.996**</b> (1.327)	<b>-3.418**</b> (1.384)	<b>-2.929**</b> (1.315)	<b>-5.159**</b> (2.041)	<b>-4.801***</b> (1.770)	<b>-3.763**</b> (1.847)
L.PTA_impsh	<b>0.344</b> (1.451)	<b>-0.199</b> (1.373)	<b>0.121</b> (1.304)	<b>-1.023</b> (1.649)	<b>1.055</b> (1.799)	<b>1.303</b> (1.539)
L.log(GDP/capita)	-0.557*** (0.151)	-0.743*** (0.157)	-0.567*** (0.148)	-0.402* (0.210)	-0.666*** (0.224)	-0.558*** (0.196)
L.DOM	-12.089** (5.779)	-17.857*** (4.694)	-71.139 (0.000)	-2.066 (1.334)	-3.304*** (1.163)	-1.396 (1.330)
L.FOR	-4.905*** (1.512)	-5.852*** (1.586)	-5.087*** (1.536)	-4.898*** (1.764)	-7.882*** (2.020)	-7.233*** (2.049)
war	1.297* (0.717)	1.250* (0.673)	0.975* (0.550)	0.755 (1.050)	1.085 (0.681)	1.409** (0.683)
L.war	-0.294 (0.692)	-0.674 (0.697)	-0.292 (0.603)	-0.645 (1.033)	-1.159* (0.651)	-1.465** (0.682)
socialist	0.296 (0.576)	0.243 (0.647)	0.215 (0.604)		-13.479*** (0.895)	-37.490*** (0.764)
africa	0.411 (0.394)	0.040 (0.367)	0.277 (0.365)	1.279** (0.634)	1.601** (0.682)	1.134** (0.571)
middleeast	0.194 (1.315)	1.024 (0.979)	0.421 (1.126)		-12.589*** (1.394)	-37.624*** (1.345)
Spain_colony	-0.496 (0.501)	-0.445 (0.476)	-0.455 (0.458)	0.738 (0.534)	0.945 (0.593)	0.445 (0.530)
UK_colony	-1.227*** (0.465)	-1.227*** (0.421)	-1.232*** (0.403)	-0.811 (0.509)	-1.160** (0.545)	-1.023** (0.509)
L.(M/GDP)	-0.003 (1.085)	0.066 (1.000)	0.106 (0.938)	-0.685 (1.585)	0.549 (1.212)	0.230 (1.191)
Duration	0.619** (0.268)			0.133** (0.064)		
Duration^2	-0.009** (0.004)			-0.003** (0.001)		
Pseudo R2	0.211		0.111	0.212		0.195
Log Lik	-232.2	-97.14	-235.4	-168.7	-63.29	-115.7
Observations	2,827	2,827	2,827	2,122	2,396	2,396

Notes: The dependent variable in regressions (1)-(3) is the enddemo measure based on the polity2 scores. The dependent variable in regressions (4)-(6) is the transition to autocracy ("tta") measure from Cheibub et al. (2010). Robust standard errors in parentheses, clustered by country in logit regressions. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

<sup>32</sup> If we run the enddemo regressions with PTA\_impsh while dropping FTA\_impsh, the estimates for PTA\_impsh remain very similar—always statistically insignificant.

It is also possible that FTAs help to sustain democracies not because of their rent destruction effects, but because democratic countries demand democracy from their FTA partners. To test for this alternative mechanism, we define another FTA import share measure, *FTA\_impsh\_moredemo*, which is constructed just as *FTA\_impsh* except that it covers only the imports from FTA partners with more democratic regimes, according to the *polity2* scores. We add that variable to the regression to capture the *additional* effect that FTAs with more democratic countries may have on the political regime. If only the pressure from more democratic partners matters for the relationship, the coefficient on *FTA\_impsh\_moredemo* would be significant but not the coefficient on *FTA\_impsh*. Table 4 shows the results. The results again indicate a sizable role of FTAs in sustaining democracy. However, there is no distinguishable additional effect from having the FTAs with more democratic partners. Thus, pressure from more democratic FTA partners is not a driver of our results.

Now, a potential problem in our analysis is the reverse causality between FTA participation and *Prob(enddemo)*. However, if our second hypothesis is valid (as we confirm to be the case in the next subsection), a greater risk of democratic failure would tend to induce *more* FTA participation. Without addressing this endogeneity issue, we find a negative effect of FTAs on *enddemo*. Had we eliminated it, the estimate would tend to be even more negative. In other words, the endogeneity problem biases the result against the hypothesis that FTAs reduce the risk of democratic failure.

But we also address the endogeneity of *FTA\_impsh* in the *enddemo* regressions formally. Although our specifications for the *enddemo* regression are all nonlinear, we now estimate it using a linear probability model (LMP), so that we can apply two-stage least squares (2SLS). Another benefit of LMP is that we can use country fixed effects in the *enddemo* regression without losing observations. To run 2SLS, we need instruments for *FTA\_impsh*. Explaining the evolution of FTA formation over time is a notoriously difficult task, as most attempts to explain FTA formation rely on time invariant or structural country characteristics. We rely instead on a recent line of research that explores “contagion” effects: when neighbors of a country engage in FTAs, the country becomes more likely to do the same. Egger and Larch (2008) and Baldwin and Jaimovich (2012), using spatial econometric methods in a bilateral data setting, find strong evidence for contagion effects in FTA formation. We therefore construct our first instrument based on the contagion effects of FTA formation. We use bilateral data to construct the contagion measure at the country level, to fit into our country-year panel data analysis. We first calculate  $FTA\_impsh_{it,j}$  for every country  $i$  that imports from country  $j$  in year  $t$  after excluding country  $i$ 's own FTA imports from  $j$ .  $Contagion_{jt}$  of country  $j$  in year  $t$  is then calculated as the weighted average of  $FTA\_impsh_{it,j}$ . We use different types of weights based on  $GDP_{it}$  and  $distance_{ij}$ , where  $distance_{ij}$  is the great circle distance between  $i$  and  $j$ .

We also use *remoteness* as instrument, defined as a country's distance to the rest of the world (ROW) weighted by countries' GDPs. As some authors (e.g. Baier and Bergstrand 2004) argue, a country's remoteness affects its propensity to engage in FTAs. Our presumption is that both *contagion* and *remoteness* affect countries' likelihood of forming FTAs but have no direct effect on the stability of their democracies.

Table 4: Enddemo regression results, **more democratic FTAs vs. less democratic FTAs**

	(1)	(2)	(3)	(4)	(5)
	logit	logit	xtlogit	Weibull	PH Cox
<b>L.FTA_impsh</b>	<b>-7.803**</b>	<b>-3.872**</b>	<b>-4.013*</b>	<b>-4.695**</b>	<b>-3.785**</b>
	(3.093)	(1.918)	(2.167)	(2.038)	(1.870)
<b>L.FTA_impsh_moredemo</b>	<b>3.901</b>	<b>1.696</b>	<b>1.779</b>	<b>2.748</b>	<b>1.739</b>
	(2.545)	(1.909)	(2.778)	(1.987)	(1.883)
L.log(GDP/capita)		-0.564***	-0.627***	-0.748***	-0.575***
		(0.152)	(0.240)	(0.157)	(0.151)
L.DOM		-12.023**	-11.877	-17.848***	-67.421
		(5.606)	(8.867)	(4.623)	(0.000)
L.FOR		-4.802***	-5.144***	-6.122***	-5.106***
		(1.587)	(1.723)	(1.650)	(1.550)
L.war		1.282*	1.293**	1.261*	0.965*
		(0.706)	(0.641)	(0.680)	(0.545)
socialist		-0.286	-0.314	-0.670	-0.285
		(0.686)	(0.679)	(0.706)	(0.601)
africa		0.227	0.350	0.217	0.165
		(0.585)	(0.707)	(0.622)	(0.584)
middleeast		0.379	0.433	0.044	0.254
		(0.404)	(0.460)	(0.356)	(0.359)
socialist_trans		0.244	0.328	0.973	0.437
		(1.194)	(1.175)	(0.913)	(1.054)
Spain_colony		-0.489	-0.476	-0.460	-0.452
		(0.487)	(0.486)	(0.473)	(0.451)
UK_colony		-1.228***	-1.306**	-1.279***	-1.238***
		(0.475)	(0.507)	(0.436)	(0.405)
L.(M/GDP)		0.023	0.017	0.151	0.155
		(1.030)	(0.885)	(0.967)	(0.905)
Duration		0.617**	0.633		
		(0.261)	(0.412)		
Duration^2		-0.009**	-0.010*		
		(0.004)	(0.006)		
Country RE			Yes		
rho			0.053		
Test rho=0 (p-value)			[0.342]		
Pseudo R2	0.073	0.212			0.112
Log Lik	-307.6	-231.9	-231.8	-96.50	-235.0
Observations	2,989	2,823	2,823	2,823	2,823

Notes: L.FTA\_impsh\_moredemo covers only the FTA partners with more democratic regimes based on Polity scores. Robust standard errors in parentheses, clustered by country in logit regressions. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 5 shows the results. Country fixed effects are included in all regressions. The first stage regression results show that *contagion* is a strong instrument for FTA\_impsh although *remoteness* is not. Together they are strong instruments as showed by the large joint F-statistics, which range from 14 to 36. An over-identification test cannot reject the null that these instruments are valid. In the second stage regressions, lagged FTA\_impsh remains significant at either the 1% or the 10% level. These results indicate that our finding that FTA participation helps to consolidate democracies is not driven by endogenous factors.

### **6.B. Are unstable democracies more likely to seek participation in FTAs?**

We now turn to our second hypothesis. We use the predicted hazard rate from the duration analysis to estimate how democratic instability affects FTA formation. The hazard is predicted based on the first logit regression in Table 3, which includes both FTA and PTA import shares.<sup>33</sup> As discussed in Section 4, to capture only the portion of the hazard unaffected by trade agreements, we predict the hazard after setting FTA and PTA import shares to zero. The Online Appendix displays a histogram of the predicted hazard, showing that it is below 15% for most of the observations in our sample.

Columns (1) and (2) of Table 6 display the regressions with the predicted hazard, where the dependent variable is the change in the FTA import share from the previous year ( $\Delta$ FTA). We also include the square of the predicted hazard to allow for possible nonlinear effects, as well as *remoteness* and *contagion*,<sup>34</sup> as defined above. Year dummies are included to capture global trends in the formation of trade agreements. To control for unobserved heterogeneity across countries, country fixed effects are also included in the regression. Since the hazard is predicted from the first stage with sampling errors, we use bootstrapping (500 replications) to adjust the standard errors.

In column (1), the coefficient of the predicted hazard rate is positive and significant at the 5% level, while its quadratic term is only marginally significant at the 15% level. This result supports our theoretical prediction that unstable democracies tend to form more FTAs than stable democracies. We can also quantify the magnitude of the effect. At the mean value of the hazard (0.025), a one standard deviation increase in the hazard (0.038) would lead to about 1.3 percentage point increase in the FTA import share. Or consider for example the case of Guatemala from 1966 to 1967, when its hazard shot up to 0.095 from 0.007. According to our estimates, this should lead to an

---

<sup>33</sup> The second-stage  $\Delta$ FTA regressions do not change in any qualitatively important way when we use the specifications in Table 2A. The same is true if we use Cheibub et al.'s (2010) definition for transition to autocracy as the dependent variable in the first stage.

<sup>34</sup> We report the results from the regressions using *contagion* with weights given by  $\log(\text{GDP}/\text{distance})$ . Results are similar with the other weights used in the definition of *contagion*.

increase of 2.3 percentage points in Guatemala's FTA import share (the actual increase was 2.65 percentage points).

Table 5: Endogeneity of FTA\_impsh in enddemo regressions, 2SLS

	IV = Contagion weighted by 1/log(dist)		IV = Contagion weighted by log(GDP)/log(dist)		IV = Contagion weighted by GDP/dist	
	(1)	(2)	(3)	(4)	(5)	(6)
	1 <sup>st</sup> stage L.FTA_impsh	2 <sup>nd</sup> stage enddemo	1 <sup>st</sup> stage L.FTA_impsh	2 <sup>nd</sup> stage enddemo	1 <sup>st</sup> stage L.FTA_impsh	2 <sup>nd</sup> stage enddemo
<b>L.FTA_impsh</b>		<b>-0.187*</b> <b>(0.113)</b>		<b>-0.349*</b> <b>(0.203)</b>		<b>-0.729***</b> <b>(0.247)</b>
lrgdppc	0.014 (0.039)	-0.059*** (0.022)	0.025 (0.041)	-0.053** (0.024)	0.021 (0.038)	-0.038 (0.026)
IDOM	0.093 (0.120)	0.076* (0.041)	0.110 (0.122)	0.100** (0.051)	0.144 (0.113)	0.157*** (0.059)
IFOR	0.064 (0.140)	-0.170*** (0.056)	0.227 (0.138)	-0.108 (0.079)	0.232 (0.159)	0.040 (0.086)
war	0.026 (0.020)	0.035 (0.027)	0.026 (0.020)	0.039 (0.027)	0.025 (0.020)	0.047* (0.028)
lwar	0.002 (0.014)	-0.018 (0.023)	-0.000 (0.014)	-0.018 (0.023)	-0.003 (0.015)	-0.020 (0.025)
L.(M/GDP)	0.080 (0.080)	-0.041 (0.044)	0.102 (0.084)	-0.019 (0.051)	0.094 (0.085)	0.031 (0.058)
Duration	-0.000 (0.003)	0.004*** (0.001)	0.000 (0.003)	0.004*** (0.001)	-0.001 (0.003)	0.004*** (0.001)
Duration^2	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)
<b>Remoteness</b>	<b>0.461</b> <b>(2.301)</b>		<b>0.357</b> <b>(2.337)</b>		<b>-0.109</b> <b>(2.260)</b>	
<b>L.contagion_ldist</b>	<b>63.304***</b> <b>(17.679)</b>					
<b>L.contagion_lgdpldist</b>			<b>0.376**</b> <b>(0.166)</b>			
<b>L.contagion_gdpdist</b>					<b>0.372**</b> <b>(0.178)</b>	
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Partial R2	0.016		0.011		0.031	
F-Statistics for IV	20.47		14.25		36.18	
Over-Identification						
Test (p-value)	0.556		0.530		0.524	
R-squared	0.363		0.350		0.353	
Observations	2,827	2,825	2,827	2,825	2,827	2,825

Notes: We use the variables contagion and remoteness as instruments for FTA\_impsh. *Remoteness* is measured as a country's distance to the rest of the world (ROW) weighted by countries' GDP. In regressions (1)-(2), *contagion* is defined as the average FTA\_impsh of all other countries (ROW) weighted by 1/log(distance) to the country in question, after we exclude this country's own FTA exports to ROW. In regressions (3)-(4), we proceed analogously but use log(GDP)/log(distance) as the weights in *contagion*. In regressions (5)-(6), we use GDP/distance as weights. Contagion variables are lagged by one year. Robust standard errors are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

In column (2), we repeat the analysis from column (1) but include DOM, FOR and duration dependence terms in the second-stage regression. The purpose of doing this is to check if our results are sensitive to these exclusion variables. The estimated coefficient of the hazard show little sensitivity to those changes.<sup>35</sup>

Table 6: Structural  $\Delta$ FTA and  $\Delta$ PTA regression results

	$\Delta$ FTA_impsh		$\Delta$ PTA_impsh	
	(1)	(2)	(3)	(4)
<b>Hazard</b>	<b>0.337**</b>	<b>0.383***</b>	<b>-0.143</b>	<b>-0.107</b>
	<b>(0.152)</b>	<b>(0.160)</b>	<b>(0.091)</b>	<b>(0.105)</b>
<b>Hazard^2</b>	<b>-0.766</b>	<b>-0.853</b>	<b>0.129</b>	<b>0.019</b>
	<b>(0.534)</b>	<b>(0.543)</b>	<b>(0.317)</b>	<b>(0.309)</b>
L.FTA_impsh	-0.132***	-0.139***		
	(0.018)	(0.020)		
L.PTA_impsh			-0.103***	-0.111***
			(0.021)	(0.024)
L.log(GDP/capita)	0.005	0.004	-0.004	-0.003
	(0.008)	(0.008)	(0.005)	(0.005)
L.DOM		0.008		-0.009
		(0.022)		(0.010)
L.FOR		0.179***		-0.032
		(0.071)		(0.034)
war	-0.010	-0.011*	0.012***	0.011***
	(0.008)	(0.008)	(0.005)	(0.005)
L.war	0.001	0.001	-0.003	-0.003
	(0.006)	(0.006)	(0.004)	(0.004)
L.(M/GDP)	0.000	0.006	0.008	0.007
	(0.020)	(0.019)	(0.010)	(0.009)
Duration		-0.000		0.000
		(0.001)		(0.000)
Duration^2		0.000		-0.000*
		(0.000)		(0.000)
remoteness	-0.334	-0.789*	-0.075	0.047
	(0.455)	(0.475)	(0.255)	(0.290)
L.contagion_lgdpldist	1.046	0.575	-0.611**	-0.548**
	(0.740)	(0.689)	(0.253)	(0.262)
Year Dummies	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Within R-squared	0.118	0.121	0.098	0.101
Observations	2,826	2,826	2,826	2,826

Notes: Hazard is the predicted hazard based on the first enddemo logit regression in Table 3, after we set L.FTA\_impsh = L.PTA\_impsh = 0. Standard errors in the second stage are corrected by bootstrapping (500 replications). \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

The last two columns of Table 6 report the results from the analogous exercise for the

<sup>35</sup> We also tried both excluding DOM and the duration dependence terms while keeping FOR in the second stage. The results remain virtually unchanged.

formation of partial-scope trade agreements,  $\Delta$ PTAs. The effect of the hazard in this case is statistically indistinguishable from zero. Thus, political instability has no identifiable effect on the formation of partial-scope trade agreements. This is consistent with the mechanism we develop, where unstable democracies seek to diminish the rents available to those in power by increasing participation in FTAs – but not in PTAs.

### 6.C. Robustness

In this section we discuss the robustness of our results along several dimensions. The accompanying tables are in the Online Appendix.

First, we investigate whether our FTA import share variable captures mostly deepening of existing FTAs or the formation of new FTAs. Although the interpretation of the results is similar in both cases, it is useful to disentangle the two effects to check whether the formation of new FTAs matters on its own. To do so, we construct an alternative FTA import share measure (*FTA\_impsh60*) by replacing current year imports with the imports *in 1960*, the first year in our sample. This measure excludes all variations in our main independent variable except those that come from the formation of new FTA links. We find that doing so increases the magnitude of the FTA coefficients slightly, relative to their levels in Table 2A.

Second, we investigate the role of the GATT/WTO in consolidating democracies. The impact of FTAs on the availability of rents is not straightforward. Due to potential trade diversion effects, it is conceivable that FTAs may lead to the creation of rents, as some authors have argued; the role of our model is to show that, once one accounts for the endogeneity of the external tariffs, FTAs actually tend to destroy rents. On the other hand, it should be relatively uncontroversial that participation in multilateral trade agreements such as the GATT/WTO, if it has any effect on the availability of protectionist rents, it must be a negative one. In light of the dynamic part of our model, this should then lower the risk of democratic breakdown.

We therefore add a WTO membership dummy as an additional regressor to our existing enddemo regressions. It turns out to be highly insignificant. Most likely this reflects the huge heterogeneity among GATT/WTO members. As Subramanian and Wei (2007) argue, there is in particular a central difference between the effects of GATT/WTO membership for countries that joined after the conclusion of the Uruguay Round (1994) and those that joined before the round, because much more stringent criteria have to be met by the acceding members since 1995 than during the GATT period. To capture this potential difference, we split the WTO membership dummy in two: *GATT\_pre1995* equals 1 from the year a country joined the GATT as a formal

member, provided that this happened before 1995, and zero otherwise; *WTO\_from1995* equals 1 from the year a country joined the WTO as a formal member, which is possible only since 1995, and zero otherwise. We find that the coefficient on *GATT\_pre1995* is always highly insignificant, while the coefficient on *WTO\_from1995* is negative and marginally significant at the 15% level. Precise estimates cannot be obtained due to the relatively small variation over time of these variables, but these results suggest that the WTO (unlike the GATT) probably has a positive effect on the consolidation of democracy in its new members. The inclusion of the two new variables notwithstanding, the effect of *FTA\_impsh* is hardly changed.

Third, motivated by the findings of Pevehouse (2005) that participation in international organizations, broadly defined, can help democracies to endure, we also test whether the effects we find for FTAs are not capturing the effect of engagement in international organizations (IOs) more generally. This would matter, as the reasoning for the latter relies on signaling effects, whereas our emphasis is on a commitment to destroy protectionist rents. Fourth, Cheibub (2006) shows that having a former military leader as the chief executive has a negative effect on democracy survival. This is confirmed by the extreme bound analysis in Gassebner et al. (2012). Thus we include the lagged number of IOs a country belongs to and a dummy variable indicating whether the executive leader is a former military officer. The estimated coefficient for formal military leader in the country is always statistically indistinguishable from zero. The coefficient on the IO variable is negative and significant in some specifications, indicating that engagement in IOs in general help democracies to endure. Nevertheless, introducing that variable has virtually no effect on our estimates of how FTA participation affects democracy survival.

Fifth, it is possible that our results are driven by a small number of observations for which alternative mechanisms may generate the same statistical association. For example, many previously socialist countries both established democratic regimes and entered in FTAs during the 1990s, but also went through many other economic and institutional changes. Are those countries alone responsible for our statistically significant coefficients? We check that by dropping all previously socialist states from our regressions. We find that the magnitude of the estimated coefficient on our FTA variable is actually somewhat higher in the restricted sample. This probably reflects the fact that some socialist countries that became democracies later reverted to autocracies, despite effective participation in FTAs (e.g. Azerbaijan).

A related concern is that some trade agreements contain rules specifying that members must be democratic. This was a pre-condition for accession of the Central and Eastern European countries to the European Union. Mercosur added a similar clause for all of its members in 1998. For those



countries, democracy and FTA participation must therefore be linked and this may have an important effect on our estimations. Again, we find that the results are very similar to those from Table 2A when we drop the observations to which such rules apply (i.e. Central and Eastern European countries once they entered the European Union and all Mercosur members since 1998).

Our democracy consolidation result could arise also because less rent-seeking democracies may be more likely to sign FTAs and may also last longer, relative to more rent-seeking democracies. This would imply that the results were driven by omission of a measure on how rent-seeking a country is. A measure for this variable is generally very hard to find, but Gawande et al. (2009) provide estimates for how much governments care about general welfare relative to special interests for 54 countries/regions. Although the subset of countries is small relative to our sample, it provides a useful proxy. Including such a measure ( $1/b$ ) in our enddemo regressions, we find that this variable is always highly insignificant, whereas the FTA\_impsh variable remains significant in all regressions despite the sharp reduction in our sample.

Another possible omitted variable that may be correlated with our measure of FTA intensity is the country's relative abundance of national resources. Ownership of natural resources in resource-rich countries tend to be concentrated in the hands of cohesive and politically strong oligarchs who benefit from international trade (and possibly also from participation in FTAs). But abundance in natural resources is also associated with higher rents and thus with the breakdown of democratic regimes (see e.g. Wantchekon 2002). This suggests a potentially *positive* bias in our enddemo regression, so not including a measure of resource abundance may artificially dampen our estimates. We confirm this presumption when we add, in turn, a measure of resource abundance constructed by Albornoz et al. (2012) and the share of ores and metal in a country's exports.

We also perform a number of robustness checks in our structural regressions.<sup>36</sup> First, does the hazard of democratic failure actually predict the formation of *new* FTAs, or just the deepening of existing FTAs? To eliminate variation in the dependent variable that is unrelated to the formation of new FTA/PTA links, we repeat the regressions in Table 6 using the measure of FTA and PTA participation that considers 1960 import shares instead of lagged ones. The estimated coefficient on the hazard falls but remains statistically significant at the 1% or the 5% level, confirming that democratic instability indeed helps to predict the formation of new FTAs (but not of new PTAs), as a more direct interpretation of our model suggests.

Second, is it really the expectation of regime change (proxied by the predicted hazard) that

---

<sup>36</sup> Because the squared term of the predicted hazard is insignificant at the 10% level in our main regressions (Table 6), we include only the level of the hazard in our robustness checks.

induces more FTAs? One alternative is that a recent change in regime, from autocracy to democracy, is what causes more FTAs. Another is that volatility in the degree of democracy has an effect on its own. Our estimated hazard may be correlated with both recent changes in political regime and the volatility of democracy. To test whether it is indeed the expectation of regime change that matters, we include in the  $\Delta$ FTA regression, in addition to the predicted hazard, one of the following variables: (1) *reg\_change*, a dummy indicating whether a country's polity2 increased from a non-positive to a strictly positive score, lagged by one year; and (2) *var(Polity)\_10yr*, the variance of polity2 scores during the last 10 years, which is used to capture recent regime instability. The first newly added variable is statistically significant but not the second. Most importantly, the estimated coefficient on the hazard remains positive and statistically significant. These results corroborate the idea that the expectation of a possible of democratic disruption matters for FTA formation.

Third, it is conceivable that our empirical results may be identifying a more general phenomenon than the model suggests. For example, governments may form more FTAs as a consequence of general political competition within an electoral system, and not specifically because of the threat from autocrats. If the risk of democratic collapse is positively associated with more political competition within a democracy, our finding that the threat of the democratic system induces FTA formation may be driven by regular political competition within a democracy. To test for that, we include a measure of political competition in the  $\Delta$ FTA regression. We use the variable POLCOMP from Polity IV, which codes the degree of political competition in the country.<sup>37</sup> Again, the newly added variable is statistically insignificant, while the predicted hazard remains statistically significant. This suggests that regular political competition in a country is not an important driver of FTA formation, unlike the expectation of democratic disruption.

Finally, negotiating and implementing an FTA is a process that can take several years. This suggests that the relationship may be stronger if we use longer lags for the explanatory variables and longer differences for  $\Delta$ FTA. We do that by running the 5-year difference in FTA\_impsh on the same explanatory variables lagged by 5 years. Alternatively, we calculate first the 5-year moving averages of the variables, before taking their 5-year lags or 5-year differences. Please refer to the notes in the appendix table for more details. When we take this longer time perspective, the estimated coefficient of the hazard remains significant and becomes considerably larger, as expected.

---

<sup>37</sup> This variable measures both the regulation of participation and the competitiveness in the political process. According to the Polity IV Dataset User's Manual, POLCOMP reaches its maximum score when "relatively stable and enduring political groups regularly compete for political influence and positions with little use of coercion; ruling groups and coalitions regularly, voluntarily transfer central power to competing groups; and no significant groups, issues, or types of conventional political action are regularly excluded from the political process."

Put together, these robustness checks boost confidence that we are capturing our intended mechanism, rather than the workings of some omitted variable that affects both FTA formation and political regimes in general.

## 7. CONCLUSION

We study the relationship between a country's participation in free trade agreements and the sustainability of its democracy. We develop a model centered on the destruction of rents caused by FTAs. The model delivers two main results. First, deeper engagement in FTAs increases the longevity of democracies. Second, political instability promotes FTA participation.

Relying on the concept of democratic capital developed by Persson and Tabellini (2009) to estimate countries' hazards out of democracy, we test and confirm our two main predictions in a sample of 116 countries over almost 50 years. Our results suggest that the rent destruction forces of FTAs constitute an important channel through which our predictions manifest. Our measure of FTA intensity is aimed at capturing those forces. Moreover, the predictions hold for "proper," GATT 1947-compatible free trade areas and customs unions, but not for partial-scope agreements based on GATT's Enabling Clause. We also find that the impact of FTAs on democracy survival is not greater when the partners are more democratic, so it is not simply that some countries demand democracy from their FTA partners. It is not that general political competition induces FTA formation either; it is the uncertainty related to the possibility of major changes in the political regime that matters.

While all this is "good news" for democratic countries involved in FTAs, we must emphasize that participation in FTAs is, unsurprisingly, no panacea: they can *help* to consolidate democracies, but their reach is limited. Our estimates make this limit clear. The effect is nevertheless comparable to that of some of the main forces highlighted in the literature, such as the spread and strength of democracies in neighboring countries. Similarly, although we have found that democratic instability compels governments to engage in FTAs, there are as well many reasons other than democratic instability that also foster participation in FTAs.

Our study pushes forward a small but growing literature (largely in political science) that seeks to uncover links between democracy and FTA participation. We depart from it by focusing on a specific mechanism—the destruction of rents in FTAs—that allows us to be more precise about the relationship. The paper also provides a clear departure from the perspective often taken by economists studying regionalism, who tend to focus on the strictly economic aspects of its causes

and consequences.<sup>38</sup> Purely economic motives certainly help, but cannot fully explain the intensity of the ongoing outbreak of free trade agreements, which have become the main trade policy instrument in many countries.<sup>39</sup> We show that the instability of democracies is another important contributing factor to this trend. Yet much remains to be known about the interplay between trade agreements and democracy. We look forward to future research to further illuminate this relationship.

## REFERENCES

- Acemoglu, Daron, and James A. Robinson (2006). *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.
- Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared (2008). "Income and Democracy." *American Economic Review* 98(3), 808-42.
- Aidt, Toke, and Martin Gassebner (2010). "Do Autocratic States Trade Less?" *World Bank Economic Review* 24(1), 38-76.
- Albornoz, Facundo, Sebastian Galiani, and Daniel Heymann (2012). "Foreign investment and expropriation under oligarchy and democracy." *Economics & Politics* 24(1), 24-46.
- Alesina, Alberto, and Guido Tabellini (1990). "A Positive Theory of Fiscal Deficits and Government Debt." *Review of Economic Studies* 57, 403-414.
- Alvarez, Michael, Jose A. Cheibub, Fernando Limongi and Adam Przeworski, A. (1996). "Classifying political regimes." *Studies in Comparative International Development*, 31(2), 3-36.
- Bagwell, Kyle, and Robert W. Staiger (2004). *The Economics of the World Trading System*. The MIT Press: Cambridge, MA.
- Bagwell, Kyle, and Robert W. Staiger (2010). "The WTO: Theory and Practice." *Annual Review of Economics* 2, 257-282.
- Baier, Scott L., and Jeffrey H. Bergstrand (2004). "Economic Determinants of Free Trade Agreements." *Journal of International Economics* 64(1), 29-63.
- Baldwin, Richard, and Dany Jaimovich (2012). "Are Free Trade Agreements Contagious?" *Journal of International Economics*, forthcoming.

---

<sup>38</sup> A notable exception is the recent study by Martin, Mayer and Thoenig (2012), who use the frequency of wars between two countries in the distant past to test whether FTAs lower the probability of war between them. They find that they do, especially when the trade gains associated with the FTA are high, indicating an interesting complementarity effect between trade and security gains from FTAs.

<sup>39</sup> A figure in the Online Appendix illustrates this trend. Indisputably, the number of free trade agreements in force has accelerated dramatically since the early 1990s. In the same figure we also show the cumulative number of transitions to democracy and transitions to autocracy throughout the world since 1948. The former trend has also intensified since the 1990s, but the opposite has happened with the latter, in line with what our results suggest.

- Barro, Robert (1999). "Determinants of Democracy." *Journal of Political Economy* 107(6), S158-S183.
- Bernhofen, Daniel, Tobias Ketterer, and Chris Milner (2012). "Preferences, rent destruction and multilateral liberalisation: the building block effect of CUSFTA." CESifo Working Paper no. 3985.
- Besley, Timothy, and Stephen Coate (2001). "Lobbying and Welfare in a Representative Democracy." *Review of Economic Studies* 68(1), 67-82.
- Calvo-Pardo, Hector, Caroline Freund, and Emanuel Ornelas (2011). "The ASEAN Free Trade Agreement: Impact on Trade Flows and External Trade Barriers." In Robert J. Barro & Jong-Wha Lee (eds.), *Costs and Benefits of Economic Integration in Asia*, Oxford: Oxford University Press.
- Cheibub, José Antonio (2006). *Presidentialism, Parliamentarism, and Democracy*. Cambridge University Press, New York.
- Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland (2010). "Democracy and dictatorship revisited." *Public Choice* 143(1-2), 67-101.
- Dixit, Avinash (2010). "Democracy, Autocracy, and Bureaucracy." *Journal of Globalization and Development* 1(1), Article 1.
- Egger, Peter, and Mario Larch (2008). "Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis." *Journal of International Economics* 76(2): 384-99.
- Estevadeordal, Antoni, Caroline Freund, and Emanuel Ornelas (2008). "Does Regionalism Affect Trade Liberalization towards Non-Members?" *Quarterly Journal of Economics* 123(4), 1531-1575.
- Frankel, Jeffery A., Ernesto Stein, and Shang-Jin Wei (1997). *Regional Trading Blocs*. Washington, DC: Institute for International Economics.
- Freund, Caroline, and Emanuel Ornelas (2010). "Regional Trade Agreements." *Annual Review of Economics* 2, 139-166.
- Gandhi, Jennifer, and Adam Przeworski (2006). "Cooperation, Cooptation, and Rebellion under Dictatorship." *Economics & Politics* 18, 1-26.
- Gassebner, Martin, Michael J. Lamla, and James Raymond Vreeland (2012). "Extreme Bounds of Democracy." *Journal of Conflict Resolution* (forthcoming).
- Gawande, Kishore, Pravin Krishna, and Marcelo Olarreaga (2009). "What Governments Maximize and Why: The View from Trade." *International Organization* 63, 491-532.
- Giavazzi, Francesco, and Guido Tabellini (2005). "Economic and Political Liberalizations." *Journal of Monetary Economics* 52, 1297-1330.
- Grossman, Gene, and Elhanan Helpman (1995). "The Politics of Free-trade Agreements." *American Economic Review* 85, 667-90.
- Grossman, Gene, and Elhanan Helpman (1996). "Electoral Competition and Special Interest Politics." *Review of Economic Studies* 63(2), 265-86.

- La Porta, Raphael, Florencio Lopez-De-Silanes, Andrei Shleifer, and Robert W. Vishny (1999). "The Quality of Government." *Journal of Law, Economics and Organization* 15, 222-79.
- Limao, Nuno (2006). "Preferential Trade Agreements as Stumbling Blocks for Multilateral Trade Liberalization: Evidence for the U.S." *American Economic Review* 96, 896-914.
- Lopez-Cordova, Ernesto, and Christopher M. Meissner (2008). "The Impact of International Trade on Democracy: A Long-Run Perspective." *World Politics* 60(4), 539-575.
- Maggi, Giovanni, and Andrés Rodríguez-Clare (1998). "The Value of Trade Agreements in the Presence of Political Pressures." *Journal of Political Economy* 106(3), 574-601.
- Maggi, Giovanni, and Andrés Rodríguez-Clare (2007). "A Political-Economy Theory of Trade Agreements." *American Economic Review* 97(4), 1374-1406.
- Mansfield, Edward D., Helen V. Milner, and Jon C. Pevehouse (2008). "Democracy, Veto Players, and the Depth of Regional Integration." *World Economy* 31(1), 67-96.
- Mansfield, Edward D., Helen V. Milner, and Peter B. Rosendorff (2000). "Free to Trade: Democracies, Autocracies, and International Trade." *American Political Science Review* 94(2), 305-321.
- Mansfield, Edward D., Helen V. Milner, and Peter B. Rosendorff (2002). "Why Democracies Cooperate More: Electoral Control and International Trade Agreements." *International Organization* 56(3), 477-513.
- Mansfield, Edward D., and Jon C. Pevehouse (2006). "Democratization and International Organizations." *International Organization* 60(1), 137-167.
- Martin Philippe, Thierry Mayer, and Mathias Thoenig (2012). "The Geography of Conflicts and Free Trade Agreements." *American Economic Journal: Macroeconomics* 4(4), 1-35.
- McLaren, John (2000). "Free Trade Agreements, Customs Unions, and the Dynamics of Political Influence." Mimeo.
- McLaren, John (2002). "A Theory of Insidious Regionalism." *Quarterly Journal of Economics* 117(2), 571-608.
- Mitra, Devashish, Dimitrios Thomakos, and Mehmet Ulubasogly (2002). "'Protection for Sale' in a Developing Country: Democracy vs. Dictatorship." *Review of Economics and Statistics* 84(3), 497-508.
- Mrazova, Monika, David Vines, and Ben Zissimos (2012). "Is the GATT/WTO's Article XXIV bad?" *Journal of International Economics* (forthcoming).
- Ornelas, Emanuel (2005a). "Rent Destruction and the Political Viability of Free Trade Agreements." *Quarterly Journal of Economics* 120(4), 1475-1506.
- Ornelas, Emanuel (2005b). "Endogenous Free Trade Agreements and the Multilateral Trading System." *Journal of International Economics* 67(2), 471-497.

O'Rourke, Kevin H., and Alan M. Taylor (2007). "Democracy and Protectionism." In Timothy J. Hatton, Kevin H. O'Rourke, and Alan M. Taylor (eds.), *The New Comparative Economic History: Essays in Honor of Jeffrey G. Williamson*. Cambridge, Mass.: MIT Press.

Persson, Torsten, and Lars Svensson (1989). "Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences." *Quarterly Journal of Economics* 104, 325-345.

Persson, Torsten, and Guido Tabellini (2009). "Democratic Capital: The Nexus of Political and Economic Change." *American Economic Journal: Macroeconomics* 1(2), 88-126.

Pevehouse, Jon C. (2005). "Democracy from Above: Regional Organizations and Democratization." Cambridge University Press.

Przeworski, Adam, Michael Alvarez, Jose A. Cheibub and Fernando Limongi (1996). "What Makes Democracies Endure?" *Journal of Democracy* 7, 39-55.

Roy, Martin (2011). "Democracy and Political Economy of Multilateral Commitments on Trade in Services." *Journal of World Trade* 45(6), 1157-1180.

Stroup, Caleb, and Benjamin Zissimos (2012). "Pampered Bureaucracy, Political Stability, and Trade Integration." Mimeo.

Subramanian Arvind, and Shang-Jin. Wei (2007). "The WTO Promotes Trade, Strongly but Unevenly." *Journal of International Economics* 72(1): 151-175.

Wantchekon, Leonard (2002). "Why do Resource Dependent Countries Have Authoritarian Governments?" *Journal of African Finance and Economic Development* 5(2), 57-77.

World Trade Organization (2011). "Preferential Trade Agreements and the WTO: from Co-existence to Coherence." The World Trade Report 2011.