WP/18/101

IMF Working Paper

The Needle in the Haystack: What Drives Labor and Product Market Reforms in Advanced Countries?

by Romain Duval, Davide Furceri and Jakob Miethe

INTERNATIONAL MONETARY FUND

 WP/18/101

IMF Working Paper

Research Department

The Needle in the Haystack:

What Drives Labor and Product market Reforms in Advanced Countries?

Prepared by Romain Duval, Davide Furceri and Jakob Miethe +

Authorized for distribution by Romain Duval

May 2018

Abstract

The political economy literature has put forward a multitude of hypotheses regarding the drivers of structural reforms, but few, if any, empirically robust findings have emerged thus far. To make progress, we draw a parallel with model uncertainty in the growth literature and provide a new version of the Bayesian averaging of maximum likelihood estimates (BAMLE) technique tailored to binary logit models. Relying on a new database of major past labor and product market reforms in advanced countries, we test a large set of variables for robust correlation with reform in each area. We find widespread support for the crisis-induces-reform hypothesis. Reforms are also more likely to happen when other countries undertake them or there is formal pressure to implement them. Other robust correlates are more specific to certain areas—for example, political factors are most relevant for job protection reforms.

JEL Classification Numbers: C52; J08; J65; L51

Keywords: Structural reforms; labor market; product market; deregulation; employment protection; unemployment benefits; Bayesian averaging of maximum likelihood estimates

Author's E-Mail Address: rduval@imf.org, dfurceri@imf.org, jmiethe@diw.de

---

[+] Romain Duval and Davide Furceri are with the IMF. Jakob Miethe is with the German Institute for Economic Research. The authors would like to thank participants to the 2017 LSE workshop on Structural Reforms and European Integration, IMF Jobs and Growth Seminar, and the 2017 Berlin Summer School of Econometrics of Panel Data and Network Analysis. In particular, we thank Nauro Campos, Steven Durlauf, Paul de Grauwe, David Sondermann, and Jeffrey Wooldridge for discussions and comments. All remaining errors are ours.

Contents Page

Tables

Figures

Annexes

I. INTRODUCTION

A prolonged period of sluggish productivity and income growth in advanced economies has led policymakers and international institutions to call for renewed efforts to undertake structural reforms, notably in labor and product markets. At the same time, structural reforms are notoriously difficult to implement, and little consensus exists over what factors can help break the deadlock, as theory is unsettled and empirical evidence is both limited and little consistent (see e.g. Acemoglu, Johnson and Robinson, 2006; Drazen, 2000; Persson and Tabellini, 2000; for labor market institutions specifically, Saint-Paul, 2000). Consider, for example, one of the most prominent hypotheses put forward in the literature, namely that crisis induces reform (e.g. Drazen and Grilli, 1993; Fernandez and Rodrik, 1991). Based on a broad cross-country time-series dataset of various macroeconomic outcomes, Drazen and Easterly (2001) find support for the hypothesis for certain types of crises but not others. Focusing on financial reforms, Abiad and Mody (2005) also find mixed effects, while Mian, Sufi and Trebbi (2014) even provide tentative evidence that crisis impedes reform. Focusing on labor and product market reforms, Hoj *et al.* (2007) find some positive effects when focusing on large changes in OECD indicators, while Agnello *et al.* (2015) do not when using Fraser Institute and IMF reform datasets instead.

This paper argues that such uncertainty reflects fundamental model uncertainty on the selection of covariates, compounded with reform measurement issues. We attempt to address both issues through a new version of Bayesian model averaging tailored to binary logit models, which we apply to a new narrative dataset of major labor and product market reforms covering 26 advanced economies over four decades.

Model uncertainty is pervasive in other areas of economics which, however, have seen gradual progress toward addressing it. Perhaps the most prominent example is the drivers of growth.[1] The key feature that has underpinned the development of model uncertainty procedures in that context (e.g. Levine and Renelt, 1992; Sala-i-Martin, 1997; Sala-i-Martin, Doppelhofer and Miller, 2004; Moral-Benito, 2012) is the existence of a wide range of growth theories without much consensus about any canonical model. This means that empirical researchers need to choose

[1] For another example in the literature on capital punishment, see Durlauf, Fu, and Navarro (2012).

amongst 2^K possible model specifications, where K denotes the (large) number of potentially relevant growth drivers. In such context, empirical results will typically be influenced by the inclusion or, more importantly, the omission of specific variables. More broadly, depending on their model selection procedure—if any, different researchers may well arrive at different conclusions even when using the same data. Model averaging alleviates such inconsistencies by comparing the robustness of regression coefficients over the entire model space.

Model uncertainty is particularly acute in the political economy of reforms, where theory is even less settled than growth theory. The list of potential reform drivers includes, amongst many others, business conditions, macroeconomic policies, structural features such as country size or demographics, external factors, or political factors such as political institutions, political capital or ideology. We summarize this literature in more detail further below. With so many potential explanatory drivers, identifying their relative importance and robustness has proven elusive. Common practice has focused on a handful of variables selected based on some expected influence on reform decisions, with little or no robustness to further controls. To make matter worse, many hypotheses are not mutually exclusive, and both included and omitted variables can be correlated.

In this paper, we address this issue by employing recently developed model averaging techniques tailored to the problem at hand. Specifically, we adapt the Bayesian averaging of maximum likelihood estimates (BAMLE) approach introduced by Moral-Benito (2012) and Dardanoni *et al.* (2015) to binary choice models explaining structural reforms. With this approach, we are able to establish and compare the (non-)robustness of a large number of potential reform drivers over the entire model space. This approach has not been applied in the political economy literature on structural reforms so far and, to the best of our knowledge, we are also the first to adapt the BAMLE methodology to binary logit models in general.

A second source of uncertainty for empirical economists trying to analyze structural reforms is reform identification. While there are widely accepted and reasonably reliable datasets on economic growth, this is not the case for structural reforms, reflecting conceptual and practical measurement issues. Early papers inferred major reforms indirectly from outcomes; for example, a collapse in inflation was supposed to indicate a significant shift in the macroeconomic policy framework (Bruno and Easterly 1996; Drazen and Easterly, 2001). Subsequent papers have typically relied on structural policy indicators produced by international organizations such as the IMF, OECD and World Bank, or by independent institutions such as the Fraser Institute (see, to give just a few examples across various reform areas, Abiad and Mody, 2005; Agnello *et al.*, 2015; Alesina, Ardagna and Trebbi, 2006; Duval, 2008; Hoj *et al.*, 2007; Wiese, 2014). These indicators attempt to measure the stance of the underlying policies by scoring and weighing their different dimensions—for example, for employment protection legislation (EPL), the length of the dismissal notice period and the possibility of reinstatement following unfair dismissal. The time-series variation of these indicators is then typically used to identify reforms, based on a specific criterion—for example, the first difference of the policy indicator in Abiad and Mody (2005), a standard deviation criterion in Duval (2008), predefined absolute changes in Agnello *et al.* (2015), or structural break tests in Wiese (2014). Each underlying indicator and approach used to identify reforms includes value judgements, including about what is a reform as opposed to a minor policy action. This is bound to further reduce the comparability of results across studies and the policy lessons that can be learned from them. More importantly, the inability to identify the exact timing of reform implementation is likely to increase measurement errors, and lead to misleading results. For example, imagine a situation when a reform is implemented in the immediate year following a crisis, while the associate structural indicator increases only few years after when the recovery takes place and growth picks up. In this situation, a researcher may erroneously conclude that reforms tend to occur during periods of stronger economic activity.

Here, we attempt to minimize value judgements and measurement error by employing a newly constructed "narrative" dataset of major reforms in four areas namely product market regulation (PMR) in network industries, EPL for regular workers, EPL for temporary workers, and unemployment benefit systems (Duval *et al.*, 2018). The main advantage of this database is that it identifies the exact timing and nature of reforms, and therefore eliminates the need for assumptions on the relation between structural reforms and regulation indicators.

Our main result supports some form of the crisis-induces-reform hypothesis across all four reform areas. High unemployment, recession and/or an open economic crisis tend to be associated with a greater likelihood of reform. The effect is economically significant. For example, an increase of 10 percentage points in unemployment (as seen in several European economies in the aftermath of the Great Recession) is associated with an increase in the probability to undertake a major EPL reform for regular contract of about 5 percentage points — that is, about twice the average probability in the sample.

We also find evidence that outside pressure increases the likelihood of reform in certain areas. Reforms are more likely when other countries also undertake them and when there is formal pressure: many product market reforms in EU countries have occurred during their accession process, and competition-relevant EU directives have also been an important factor behind deregulation.

In addition, while there is generally little robust evidence of an important role of political factors in driving reforms, EPL for regular workers stands out as an important exception and tend to occur in right-leaning governments; this is consistent with theories that highlight the ability of entrenched interests to block structural reforms (e.g. Tommasi and Velasco, 1996). Lastly, we have an interesting list of non-robust variables. In particular, the political business cycle seems to have less importance than commonly assumed, and, with the exception mentioned above, we do not find any evidence for an ideological bias—there is no robust difference between left- and right-of-center governments' propensity to undertake reform.

This paper is related to two strands of literature—the political economy of reforms, and Bayesian model averaging. The extensive literature on the political economy of reforms has relied on single-model specifications and failed to tackle the fundamental issue of model uncertainty. Together with issues in identifying reforms, this has hindered the emergence of consensus regarding the drivers and non-drivers of reforms. Our study is the first to seriously tackle model uncertainty in this strand of the literature, using a new dataset that readily identifies major labor and product market reforms based on a narrative approach. In doing so, we are able to synthesize and compare existing hypotheses regarding reform drivers. We also contribute to the model averaging literature by providing an application of Bayesian averaging of maximum likelihood estimators to binary logit models. To our knowledge, our paper is the first to do so.

The remainder of this paper proceeds as follows: Section 2 introduces the model uncertainty problem in determining the drivers of structural reform by giving a short synthesis of the literature. Section 3 introduces our database of major labor and product market reforms as well as the other data sources used in the exercise. Section 4 outlines the BAMLE methodology as well as the adjustments we carry out to apply it to binary logit models. Section 5 presents the main results as well as extensive robustness checks. Section 6 concludes.

II. THE MODEL UNCERTAINTY PROBLEM IN THE POLITICAL ECONOMY OF STRUCTURAL REFORM

At the risk of over-simplification, the theoretical and empirical literature on the political economy of reforms has emphasized six broad categories of potential drivers of reforms:

- business conditions, with particular emphasis on the role of crises;
- macroeconomic (monetary and fiscal) policies;
- structural features of the domestic economy (e.g. size, trade openness, demographics, income inequality; stringency of existing regulations, and thereby need and scope for reform);
- external factors, including formal and informal international pressure on the domestic economy to undertake reform;
- political factors, including ideology, structural features of the political system, and conjunctural features (e.g. political cycles and crises);
- reform packaging, sequencing and momentum, and more broadly the role of reform strategies in overcoming the resistance to reform that stems from the various factors listed above.

We review these six categories below, with the view to providing a sense of the wide range of possible political economy drivers of reform and the lingering uncertainties regarding their relative importance, rather than a comprehensive survey of the literature.

The large welfare costs of economic or financial crisis can break the deadlock over welfare-enhancing measures that could not be adopted otherwise due to conflict over their distributional consequences (Drazen and Grilli, 1993). Crisis can also reveal information, and thereby raise awareness, about the unsustainability of current policy arrangements and the need for change (Tommasi and Velasco, 1996). Almost by definition, drastic changes in policies and institutions should take place when current settings are no longer tenable, that is, when they result in crisis (Rodrik, 1996). These theoretical arguments suggest a non-linear impact of business conditions on reform adoption—acute crisis matters disproportionately more than just bad economic conditions. At the same time, the empirical literature has not provided unequivocal support to the "crisis-induces-reform" hypothesis, partly reflecting differences across studies in the nature and definition of the crises and structural reforms considered. Early literature, partly based on case studies, has supported the view that crisis, hyperinflation and major fiscal and external imbalances lift obstacles to reform that would otherwise prevail in normal times (Nelson, 1990; Grindle and Thomas, 1991; Haggard and Kaufman, 1992; Bates and Krueger, 1993; Haggard and Webb, 1994; Williamson and Haggard, 1994). Drazen and Easterly (2001) investigate the hypothesis more systematically and confirm it for certain types of crises (e.g. very high inflation) but not others (e.g. negative growth). Pitlik and Wirth (2003) and Agnello *et al.* (2015) find that crisis strengthens governments' economic liberalization efforts in most areas, as do Lora and Oliveira (2004) for Latin America. Consistently, Campos and Horvath (2012b) find that good economic conditions facilitate reform reversals. Abiad and Mody (2005) find mixed effects on financial reforms, varying depending on the nature of the crisis. Mian, Sufi and Trebbi (2014) even provide tentative evidence of an adverse effect, for which they find an empirical explanation in the increased polarization and fractionalization of the population after financial crises. For labor and product market reforms, previous studies have tended to yield mixed results: using OECD indicators, Hoj *et al.* (2007) find some positive effects of crisis on the likelihood of liberalization while, using Fraser Institute (Economic Freedom of the World) and IMF data, Agnello *et al.* (2015) do not.

Accommodative macroeconomic policies may enhance the likelihood of reforms through three channels. First, fiscal policy can be used to compensate reform losers, which should make it easier to overcome the status quo that typically results when there is a number of veto players whose support is critical (e.g. Fernandez and Rodrik, 1991). Second, and related, monetary and fiscal policy may facilitate reform by helping bring forward the long-term gains of reforms that may otherwise entail short-term costs (for empirical evidence on this mechanism, see Duval and Furceri, 2018). Third, fiscal consolidation erodes governments' political capital, hindering their ability to carry out structural reforms (e.g. Eichengreen *et al*., 1998). Duval (2008) finds empirical support for a positive effect of sound fiscal positions and accommodative fiscal policy on the likelihood of labor and product market reforms for a panel of OECD countries. Similar arguments have been made regarding the role of accommodative monetary policy (Draghi, 2016) and, more broadly, of independent monetary policy and therefore of a flexible exchange rate regime (Duval and Elmeskov, 2005). Running against this is the "TINA" (there is no alternative) argument, namely that there is no alternative to undertaking structural reforms to enhance labor and product market flexibility when monetary policy autonomy is lost (e.g. Bean, 1998).

Structural features of the economy have also been put forward as influential forces. Small open economies may be more amenable to reform due to greater exposure to competitive pressures and international policy diffusion (e.g. Belloc and Nicita, 2011). Demographic aging could make it harder to pass pension reforms (e.g. Galasso, 2006) but possibly easier to pass labor market reforms that affect only the working-age population, or product market reforms—such as opening services sectors to competition— that benefit all consumers through lower prices but not necessarily all workers. High inequality, by exacerbating distributional conflict, can produce anti-growth tax and regulatory policies (Persson and Tabellini, 1994). The stringency of regulation itself has ambiguous effects on the likelihood of reforms. On the one hand, a country with greater scope for reform should be more likely to exploit that scope. On the other hand, since most reforms create clear and immediate reform losers but spread benefits to larger parts of society in a more distributed and long-term fashion, a high level of regulation can find fierce and successful defenders (Olson, 1965; Tommasi and Velasco, 1996; Giuliano, Mishra and Spilimbergo 2013). In the labor market area, tight labor market regulation can create its own constituency—particularly low-productivity workers who would otherwise lose their job—strengthening opposition to reform and the persistence of rigidities (Saint-Paul, 1997). Perhaps reflecting this ambiguity, previous empirical studies have not found clear-cut effects of the initial stance on the likelihood of labor and product market reforms (e.g. IMF, 2004).

International pressure to reform can stem from both formal and informal forces. Some formal institutional arrangements incentivize or even mandate reforms. A prominent one in our context is European Union membership, which likely fostered reforms as a result of both the accession process and European Commission directives, for example in the area of network industry deregulation. The existence and actions of such institutions may partly be interpreted as a formalization of broader informal forces, however. Informal international pressure can result from learning about the experience of reforming and non-reforming country peers, which changes beliefs about the consequences of policy action (e.g. Krueger, 1993; Tommasi and Velasco, 1996).[2] Learning models have been built and tested in several studies, with broadly positive albeit somewhat mixed results. Meseguer (2006) finds supportive empirical evidence in the areas of trade liberalization and privatization, but not in others, over a broad panel of countries. Gassebner, Gaston, Lamla (2011) find that broad economic liberalization spills over most strongly across culturally and geographically close countries, and so do Fidrmuc and Karaja (2013) for Eastern European and former Soviet Union countries in the 1990s. IMF (2004) also notes that reforms may be more clustered across countries in the context of multilateral and regional integration efforts. In particular, regional leaders may set the benchmark of liberalization policies which are subsequently adopted by regional followers (see e.g., in the context of financial reform, Elhorst, Zandberg and De Haan, 2013).

Two broad sets of political drivers of reform have been highlighted in the literature. One set relates to the political cycle, and emphasizes political capital requirements to break reform deadlock as well as re-election pressures. Because reforms, including of labor and product markets, may entail short-term costs while gains can take time to materialize, reform should be less likely before elections and more likely in the beginning of a term (e.g. Alesina, Ardagna, and Trebbi, 2006; Bonfiglioli, and Gancia, 2013). Other timing effects mentioned in the literature concern the power of veto players and the strength of the opposition (Martinelli and Tommasi, 1995), which is likely to be weaker early in the term when the government's political capital to spend on reform is highest. An extreme version of the political cycle argument is the role of political crisis. Focusing on trade and labor market liberalization over a historical sample of about 100 economies, Campos, Hsiao, and Nugent (2010) find political crises to be more influential than economic ones.

The other set of political factors is unrelated to the political timing of reform, and instead focuses on features such as the fractionalization of parliament or the government coalition, the political orientation of the government or other intrinsic characteristics of the political system. Conflicting results have been found across the literature in this area. For example, while in theory

[2] For example, a widespread explanation for the emergence of the so-called "Washington Consensus" (Williamson, 1990; Williamson and Haggard, 1994; Rodrik, 1996), which inspired economic liberalization in many emerging and developing economies in the late 1980s and 1990s, has been the failure of alternative policies (e.g. import substitution) in previous decades.

fractionalization in the government coalition should increase the ability of small parties to block reforms (Alesina and Drazen, 1991), its impact has ranged from entirely insignificant (Wiese, 2014) to highly significant (Bortolotti and Pinotti, 2008; Alesina, Ardagna, and Trebbi, 2006) in empirical studies. Likewise, as noted for example by Galasso (2014), it has proven difficult to confirm empirically the so-called partisan bias that highlights the pro-market orientation of conservative parties and the resistance of leftwing parties, alongside those of their electorate, to carry out structural reform (e.g. Bortolotti, Fantini, and Siniscalco, 2003, regarding privatizations). Two interpretations have been put forward. One is the "it-takes-a-Nixon-to-go-to-China" hypothesis, which holds that the capability of governments to convince their electorate of the need for reforms will be greater especially if those run against their ideological predisposition (Cukierman and Tommasi, 1998). In the context of labor and product market reforms, while a reforming right-of-center government may face the combined resistance of the leftwing electorate, trade unions and other civil society groups, a left-of-center government will be less likely to be accused of pushing through reforms on ideological grounds and may therefore be more likely to succeed. Another explanation attributes the unsettled empirical evidence to a gradual shift in 'leftwing' governments' ideology away from socialism toward social democracy or "social liberalism" in recent decades (Potrafke, 2009).

Reform strategies, such as packaging or sequencing reforms, may help overcome some of the political economy obstacles discussed above. In our context, bunching together reforms that lower real wages or entail transitory macroeconomic costs—such as relaxing EPL—with others that raise real wages—such as lowering entry barriers in product markets—may increase the chance of reform adoption (e.g. Cacciatore *et al.*, 2016). As for sequencing, it has been suggested that deregulating product markets first would lower monopoly rents, making it easier for workers to accept subsequent reform of labor market institutions that were designed to capture those rents (Blanchard and Giavazzi, 2003).

The learning argument in favor of cross-border spillovers within one reform area can alternatively apply across areas within one country (Volden, Ting, and Carpenter 2008): reform may generate its own momentum, leading to 'reform cascades' (e.g. Tommasi and Velasco, 1996). For example, focusing on the transition of former Soviet Union countries, Golinelli and Rovelli

(2013) find that successful reforms that improved economic performance generated support for more extensive reforms later on, and vice versa.

Our empirical analysis features variables that capture all these categories of reform drivers, but ignores a few other potential forces, due to our focus on advanced economies. We do not cover the role of conditionality under adjustment programs (see e.g. Smets, Knack and Molenaers, 2012, for an analysis of World Bank conditionality) as the latter have been rare events in our sample. Likewise, while some studies have documented a positive effect of democracy on structural reforms (e.g. Giuliano, Mishra and Spilimbergo, 2013), we do not test for it here because the vast majority of OECD countries has long scored very high on this dimension in the often-used Polity IV index. Partly related, we do not do justice to some specifics of the literature on reforms in transition economies, given that our sample incorporates only one of them. For example, it has been argued that transition economies carried out reforms regardless of the color of government (e.g. Roberts and Saeed, 2012) or democratization (Campos and Horvath, 2012a). These peculiarities are ignored here. Our goal is to shed light on the relative importance of the various available hypotheses and their robustness to model specification.

III. REFORM IDENTIFICATION AND DATA

3.1 Employing a narrative database of structural reforms

We attempt to minimize value judgements and measurement error in the identification of reforms by employing a newly constructed "narrative" dataset of major reforms in four areas namely PMR in network industries, EPL for regular workers, EPL for temporary workers, and unemployment benefit systems (for full details, see Duval *et al.*, 2018).

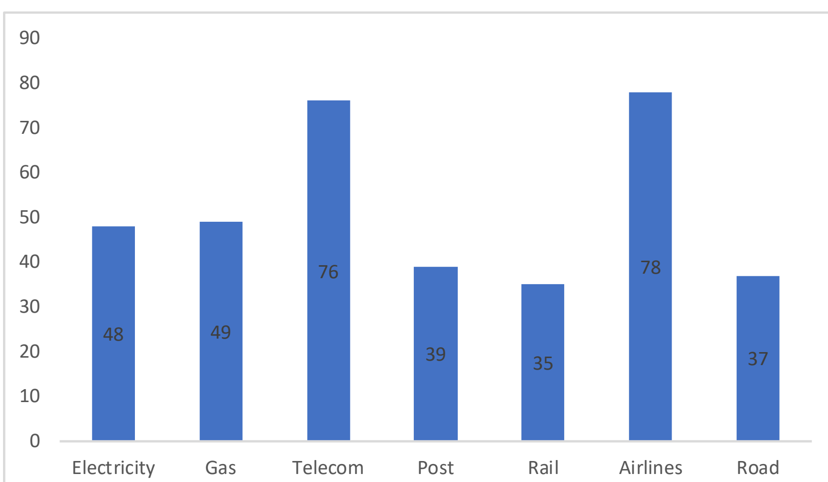
In the spirit of Romer and Romer (2010, 2017), this "narrative" approach identifies major reforms (deregulation measures), including their exact nature and timing, on the basis of simple pre-determined criteria, including the language used by the OECD to describe the policy change in its *OECD Country Survey*—the regular country surveys published by the OECD—for the country and year considered. Specifically, first, for each of 26 advanced countries and each year over 1970-

2013, all legislative and regulatory actions mentioned in all past *OECD Economic Surveys* are recorded. Second, among all those actions, major measures are identified as those that meet at least one of three alternative criteria: (i) a narrative criterion based on OECD staff's judgement on the significance of the reform at the time of adoption; (ii) whether the reform is mentioned again in subsequent Economic Surveys, as opposed to only once when the measured is adopted; (iii) the magnitude of the change in the corresponding OECD indicator, when available. The timing of each reform is precisely identified, and its content carefully documented. This approach also eliminates the need for assumptions on the relation between structural reforms and regulation indicators.

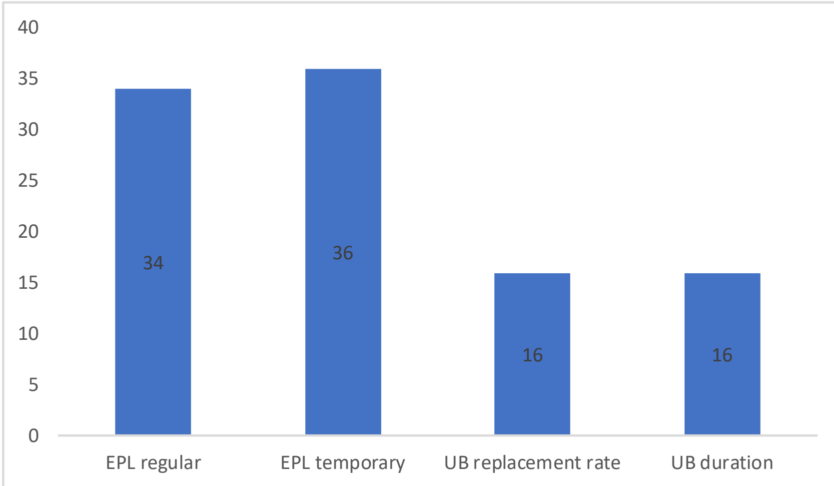
Figures 1-3 present stylized facts on reforms—focusing only, as this paper does, on decreases in regulation and leaving out increases. Major liberalizing reforms appear to have been more frequent in product markets than in labor markets in the last decades. Figures 1a and 1b, which provide the total number of reforms identified in the sample, illustrate this heterogeneity of reform efforts across regulatory areas. In product markets, major reforms have been most frequent in telecoms and airlines. As regards labor markets, major changes in EPL have been more common than major changes in unemployment benefit systems (Figure 1, Panels A and B). It is also worth noting that reform reversals are extremely rare events in practice, so we ignore this issue here.

Figure 1. Number of Major Reforms (26 advanced economies, 1970-2013)

Panel A. Product Market Regulation



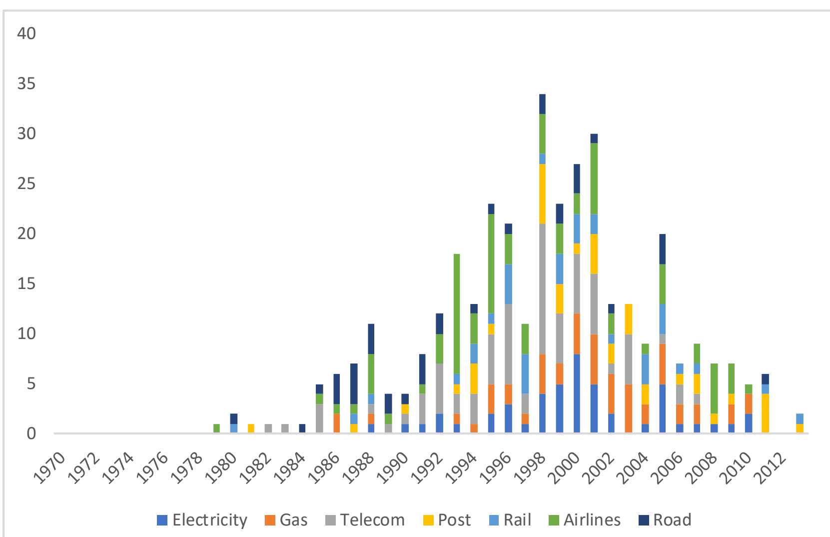
Panel B. Labor Market Regulation



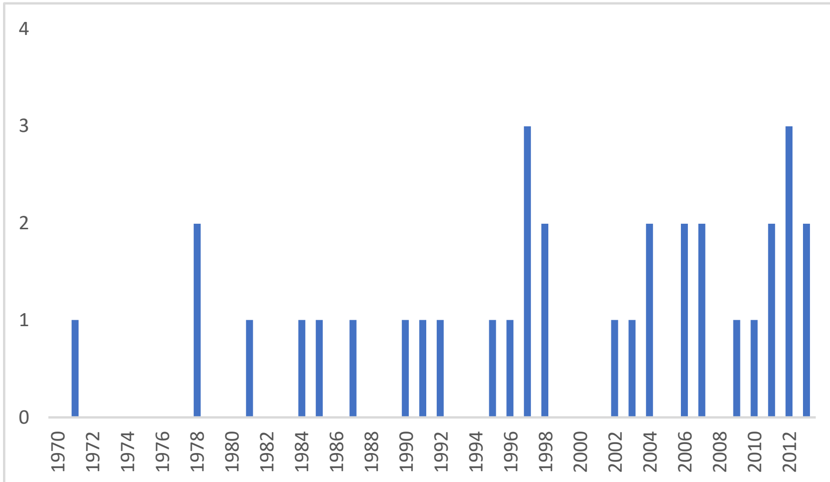
Liberalizing reforms have been predominantly implemented during the 1990s and the 2000s (Figure 2, Panels A through D). This is most striking for product market reforms, which were clustered around the late 1990s and early 2000s, partly reflecting the EU-driven liberalization process in European countries over this period. In labor markets, gradual liberalization took place starting from the 1980s. This pattern holds true for both unemployment benefit systems and EPL.

Figure 2. Distribution of Major Reforms Across Time (26 advanced economies)

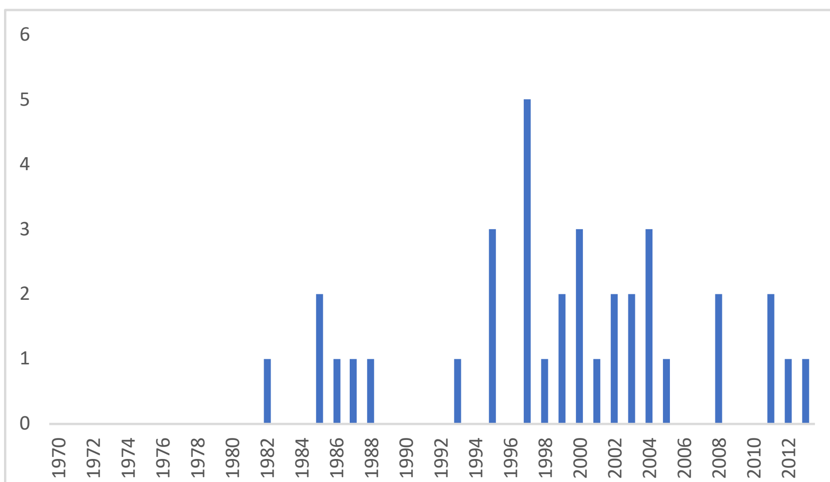
Panel A. Product Market Regulation



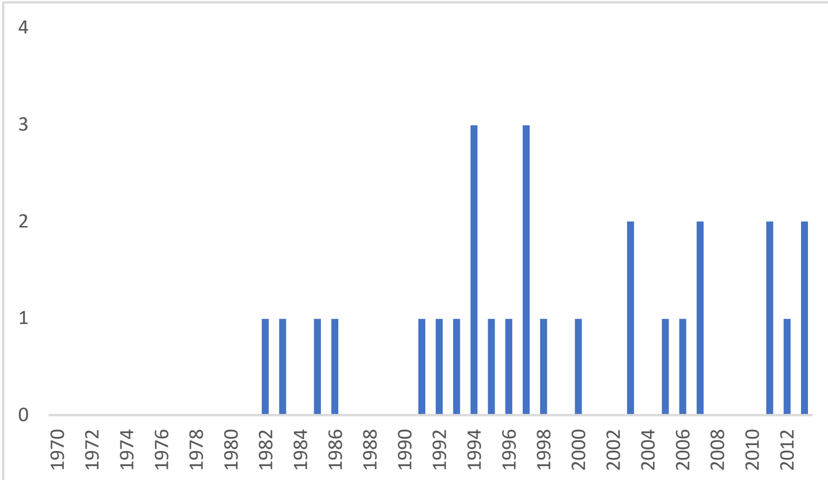
Panel B. Labor Market Regulation: Regular Contracts



Panel C. Labor Market Regulation: Temporary Contracts



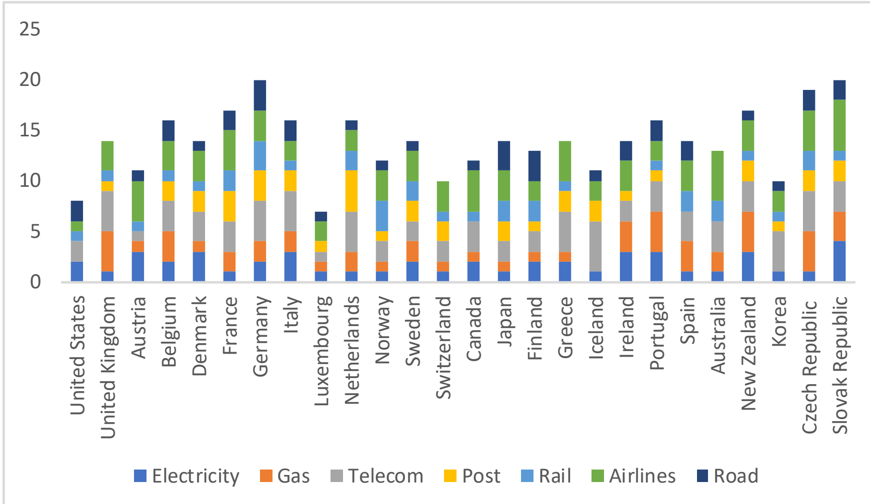
Panel D. Unemployment Benefits



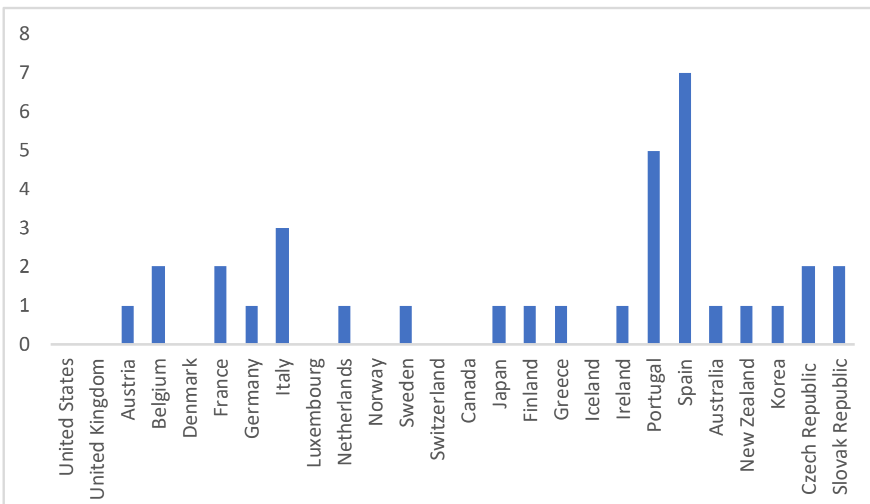
In terms of geographical distribution, EU countries took more actions than non-EU countries on average, reflecting to a large extent the greater scope for action in the former group (Figure 3, Panels A through D). While in product markets the frequency of reforms was generally similar across country groups, in labor markets southern European countries (e.g. Portugal, Spain) took many more significant actions, particularly towards easing EPL for both regular and temporary workers (Figure 3, Panels B and C). Concerning unemployment benefit systems, several countries increased or maintained the generosity of their systems during the 1970s and early part of the 1980s before reducing it later on (Figure 3, Panel D). Reforms touched roughly equally on replacement rates and duration.

Figure 3. Number of Major Reforms by Country (1970-2013)

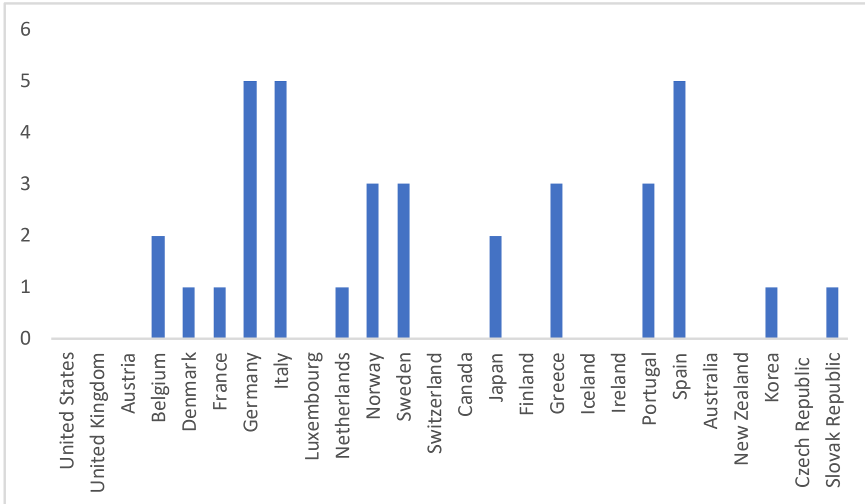
Panel A. Product Market Regulation



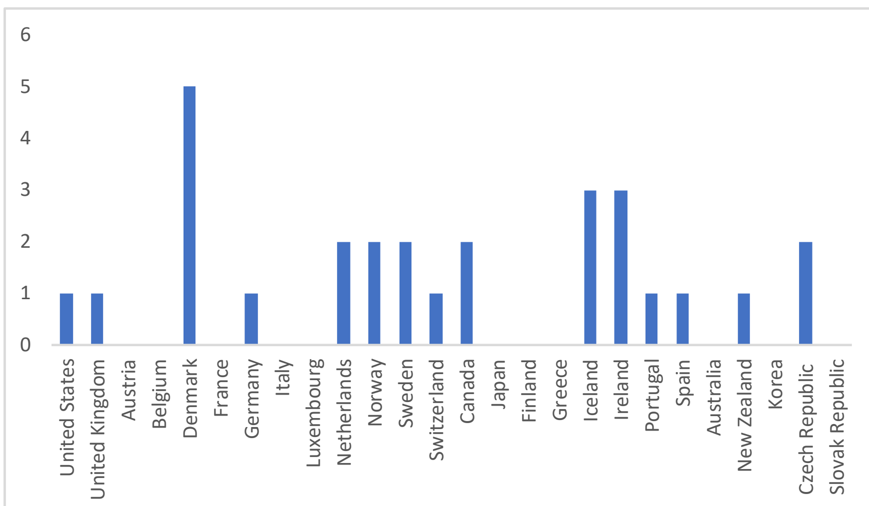
Panel B. Labor Market Regulation: Regular Contracts



Panel C. Labor Market Regulation: Temporary Contracts



Panel D. Unemployment Benefits



3.2 Other data sources

Since we aim to provide a synthesis and comparison of commonly cited studies in the literature on structural reforms, we use the most commonly employed data sources wherever available. Table 1 presents descriptive statistics for all variables used in the analysis. Their construction and sources are briefly reviewed below. The lags employed rule out some simultaneity concerns without claiming to achieve causal inference. Instead, we establish the relative robustness of different covariates that are claimed to drive reforms in the studies outlined above.

Table 1. Descriptive Statistics

	variable	construction	obs.	mean	st.dev	min	max	source
Reform database	product market reforms	narrative database	841	0.05	0.22	0	1	Duval, Furceri, Hu, Jalles, and Nguyen (2018)
	labor market reforms (regular contracts)	narrative database	841	0.04	0.19	0	1	
	labor market reforms (temporary contracts)	narrative database	841	0.04	0.20	0	1	
	unemployment benefit reforms	narrative database	836*	0.03	0.18	0	1	
Business conditions	GDP growth	growth rate	841	2.41	2.67	-9.28	12.44	IMF WEO
	unemployment	lag(rate)	841	6.63	3.66	0.18	27.47	IMF WEO
	deep recession	count variable: years in lowest 20% of GDP growth	841	0.32	0.71	0	5	based on GDP growth
	crisis	dummy (Bank, Currency and/or Sov. debt crisis/restructuring)	841	0.11	0.31	0	1	Valencia and Laeven (2012)
Macroeconomic policies	exchange rate regime	1(loose) : 15(tight)	841	7.30	4.37	1	14	Ilzetzki, Reinhart, and Rogoff (2017)
	taylor rate	lag(residual of regressing 3 months interest rate on the cpi rate and cyclical gdp)	841	0.09	3.19	-17.23	8.91	based on IMF data
Structural features of the economy	openness	lag((exp + imp)/GDP)	841	0.76	0.61	0.11	4.43	based on IMF Data
	old age dependency	Ratio	841	14.18	2.97	4.13	25.01	World Bank
	gini coefficient (net)	based on net income	841	28.41	4.07	19.70	37.80	IMF
	gini coefficient (market)	based on gross income	841	44.63	4.81	32.10	56.60	IMF
	government debt	in % of GDP	841	0.58	0.341	0.023	2.38	IMF
	reg. index: product markets	lag(level); *used for PMR*	841*	3.66	1.42	0.79	5.98	OECD
	reg. index: labor (regular contracts)	lag(level); *used for EPL regular*	685*	2.14	0.93	0.26	5	OECD
	reg. index: labor (temporary contracts)	lag(level); *used for EPL temporary*	685*	1.86	1.34	0.25	5.38	OECD
	reg. index: gross replacement rate	lag(level); *used for unempl. benefits*	799*	28.64	13.27	0.35	64.94	OECD

	over 5-year spell							
External factors	EU accession	dummy based on accession year (t-3 : t-1); *excluded for EPL regular*	841	0.02	0.15	0	1	manually constructed
	EU directives	count (t-3 : t-1); *used for PMR*	841	0.80	1.05	0	3	Bouis, Duval, and Eugster (2016)
	int. spillovers	adds reforms in same are in other countries (t-3 : t-1)						Duval, Furceri, Jalles, and Nguyen (2018)
Political factors including ideology, structural features of the political system and conjunctural features	exec r.l.c (cont)	right: 1, center: 2, left: 3	841	1.92	0.91	1	3	DPI
	exec left	Dummy	841	0.37	0.48	0	1	DPI
	union density	% of employees with right to bargain	841	38.34	21.77	7.54	99.07	OECD
	centralization gov.	sum of squared seat share of all gvt parties	841	0.71	0.27	0.18	1	DPI
	vote share gov.	in %	841	47.70	13.32	0	84.42	DPI
	gov. controls all houses	Dummy	841	0.21	0.408	0	1	DPI
	months to election	count of months left	841	23.55	13.79	1	58	Gupta, Liu, and Mulas-Granados (2016)
	elections next year	Dummy	841	0.29	0.45	0	1	based on months to elections
	years left in term	Count	841	1.68	1.241	0	4	DPI
	years executive spent in office	Count	841	3.86	2.91	1	18	DPI
momentum	domestic packaging	adds reforms in same country but other areas (t-3 : t-1)						Duval, Furceri, Hu, Jalles and Nguyen (2018)

*: in the four BAMLE exercises, the sample is limited by and therefore reduced to the available data for the relevant regulation indicator which is included in every model to provide minimal structure. The results for product market reforms are based on 841 observations, those for reforms of employment protection legislation on 685 observations and those for unemployment benefit reforms on 799 observations. All variables are then reduced to these dimensions.

Business conditions and economic crises. Variables on weak economic conditions and recession are taken from several sources. The growth rate of GDP, our main business cycle variable, is taken from the IMF World Economic Outlook database and enters as the actual annual rate of real GDP growth in the current year. The unemployment rate enters with a lag and is taken from IMF and OECD data. A deep recession variable counts the years in which GDP growth is in its lowest 20% across our panel dataset; similar results are obtained for more stringent cut offs (10% and 15%). We also use a crisis variable that takes value 1 if there was either a currency crisis, a sovereign debt crisis, or a sovereign debt restructuring in the last three years, based on data from Valencia and Laeven (2012).

Macroeconomic policies. To explore the impact of exchange rates, we employ a continuous variable on exchange rate regimes provided by Ilzetzki, Reinhart, and Rogoff (2017). We also build a variable capturing the monetary policy stance (taylor rate). This variable is constructed as the residual from a regression of the short-term interest rate on the CPI inflation rate and the cyclical component of GDP (HP detrended) with all variables taken from the IMF World Economic Outlook database.

Structural features of the domestic economy. We employ a number of variables to capture economic and structural conditions: (i) the first lag of trade openness—measured as the ratio of exports and imports to GDP (data taken from IMF World Economic Outlook database); (ii) the old age dependency ratio, defined as the ratio of the population above 65 and above to the working age population, in percent (taken from the World Bank Development Indicators database); (iii) two variables that capture inequality, namely the Gini coefficients based on net and gross income taken from the Standardized World Income Inequality Database; and (iv) government debt as a percent of GDP (sourced form IMF World Economic Outlook database). To capture reform pressure when existing regulation is stringent, we use (both the lag and initial sample value of) OECD indicators of the policy stance in each of our four reform areas (summary indicators of PMR, EPL for permanent and temporary contracts, and gross replacement rate of unemployment benefits over a five-year unemployment spell). The first lag of these indicators is the one and only variable we include in every specification for each reform and only one indicator is included for each reform as indicated in italics in Table 1.

External factors. An EMU accession dummy (again acting in t-1:t-3) is constructed based on the date of accession to the monetary union. This dummy variable does not coincide with a single reform in EPL for regular contracts which is why we omit it from the analysis of that specific reform area. It is included for the other three. We also use data on EU directives in product markets, taken from Bouis, Duval, and Eugster (2016), who in turn construct them using the comprehensive historical list of all single market directives available on the European Commission website; specifically, we include the number of directives from t-3 to t-1. This variable is only used in the variable list for product market reforms. To capture foreign ("peer") pressure for reform on the domestic economy, we collect reforms in the same field in all other countries in the sample. We compute and use the total number of these foreign reforms over the last three years (t-1:t-3).

Political factors. Data on the political orientation of the government enters in two forms. First, we introduce a continuous variable measuring ideology from right (0) to left (2) through center (1). In order to investigate the effect of leftwing governments, which has received some attention in the literature, we also include a dummy capturing such governments. These data are based on the Database of Political Institutions. Union density is the share of workforce that is a member of a union, taken form the OECD. The centralization of the government (measured as the squared parliamentary seat share of all government parties in office), the vote share of the government, a dummy variable indicating if it controls all houses, the number of years a government has already spent in office, and the number of years it has left in its term, are all taken from the Database of Political Institutions. Detailed data on the months left until the next legislative election is taken from Gupta, Liu, and Mulas-Granados (2016), extended manually where necessary. We construct a dummy variable indicating forthcoming elections, which takes value 1 if the next election is scheduled to take place in less than 12 months.

Domestic packaging. Finally, we consider as an explanatory variable the occurrence of reforms in the other 3 fields in the same country. This aims to test whether reforms are more likely to be introduced as a package, as this can facilitate their implementation. Since it is not clear how the reforms are sequenced over the different areas, we use a window from t-2 : t+2; results are robust to considering instead a window from t-1 : t+1.

IV. METHODOLOGY

Our brief overview of the literature shows that researchers have not been able to find consensus on the main drivers of structural reform. One reason for this is fundamental model uncertainty concerning the choice of covariates, a pervasive issue in the absence of any dominant model to explain structural reforms in each of the four areas analyzed. This motivates the use of a model averaging estimator that explicitly takes model uncertainty into account and compares the robustness of regression coefficients over the entire model space (Leamer, 1978; 1985; Koop, 2003). While model uncertainty exists along several dimensions (interaction terms, lag choices, or functional forms) we focus on the selection of covariates as this emerges as the main problem identified in the literature review above. The method we employ is thus aimed at checking the robustness of potentially relevant variables across a plethora of different models without being dependent on having found the one true model, if it even exists.[3] It should be noted that the BAMLE approach does not provide a causality tests; by using the lag structures shown in Table 1 in a yearly dataset, we circumvent this issue somewhat but we leave it to future research to provide causality tests.

In this section, we introduce the relevant statistical theory in our context: Bayesian averaging of maximum likelihood estimates or BAMLE (see Moral-Benito, 2012: Dardanoni *et al.*, 2015). We then describe our choice of priors as well as sampling choices. Lastly, we introduce a tailoring of BAMLE to the study of binary choice logit models.

4.1 Bayesian Averaging of Maximum Likelihood Estimates

For expositional purposes, consider a linear specification $y^* = x\beta + e$ where y^* is a latent variable we attempt to relate to a (large) set of variables in x and e is a vector of random shocks. Without fixing the dimension of x and without choosing covariates in *ad hoc* fashion, there are 2^K

[3] The point has been raised that asymptotically, Bayesian model averaging places all posterior weight on a single model (see Li and Dunson, 2017, for a related working paper). We do not address that issue here but follow the "broad pragmatic view […] that posterior model probabilities provide a useful basis for inference and prediction under model uncertainty" (ibid, p. 2). The standard algorithm we use converges on a region of model sizes, never on a single model.

possible models M_j with $j = 1, 2, \ldots, 2^{\mathrm{K}}$. Each model including a specific set of variables x^j will produce a set of parameters β^j. Therefore, as we move through the model space, we face different parameter estimates for each variable considered. In a Bayesian logic, this is reasonable as parameters are not assumed to have one true value but are random variables. Bayesian model averaging divides this parameter space into different regions of which model M_j is one (Sala-i-Martin, Doppelhofer, and Miller, 2004; Hjort and Claeskens, 2008; see Hjort and Claeskens, 2003, for frequentist comparisons). Thus, there are two sources of uncertainty: the parameters and the model. In Bayesian terms, we define a model by a likelihood and a prior density (Koop, 2003). The posterior density of the parameters of model M_j with which we want to explain the data D is given by:

$$g\left(\beta^j \middle| D, M_j\right) = \frac{f\left(D \middle| \beta^j, M_j\right) g\left(\beta^j | M_j\right)}{f\left(D | M_j\right)} \tag{1}$$

Where $f(\mathrm{D}|\beta^j M_j)$ is the likelihood function and $g(\beta_j|M_j)$ is the prior on the parameters for model M_j. On top of the uncertainty surrounding β^j, there is uncertainty concerning model M_j and we need the data to inform us about the likelihood of that model being the true model. The researcher has some prior belief regarding this likelihood denoted by $P(M_j)$. This prior that M_j is the true model can be updated by being informed by summarizing the information about M_j in the data via the integrated likelihood of model M_j denoted by $f(\mathrm{D}|M_j)$. The resulting posterior *model* probability $P\left(M_j \middle| D\right)$ can be represented using Bayes rule as:

$$P\left(M_j \middle| D\right) = \frac{f\left(D \middle| M_j\right) P\left(M_j\right)}{f(D)} \tag{2}$$

Equation 2 thus gives the posterior probability of model M_j being the true model depending on the prior $P(M_j)$ and the information about M_j in the data summarized in $f(\mathrm{D}|M_j)$. The prior $P(M_j)$ does not involve the data while the integrated likelihood follows from equation 1 (see Moral-Benito, 2012; Raftery, 1995 for details). Summing these posterior model probabilities $P\left(M_j \middle| \mathrm{D}\right)$ over the whole model space, with β a function of the β^js, one can show that the posterior of the parameters taking into account model uncertainty is a weighted sum of the posterior density of the parameters in model M_j weighted by the posterior model probability of M_j over the entire model space:

(3) $$g(\beta|D) = \sum_{j=1}^{2^k} P(M_j) \left[\frac{f(D|M_j)P(M_j)}{f(D)}\right]$$

$$= \sum_{j=1}^{2^k} P(M_j|D) \left[\frac{f(D|\beta)g(\beta|M_j)}{f(D|M_j)}\right]$$

$$= \sum_{j=1}^{2^k} P(M_j|D) g(\beta|D, M_j)$$

This shows that the posterior density of all parameters is the sum of the posterior distribution of β under each model M_j weighted by the posterior probability of that model. Here, following Sala-i-Martin, Doppelhofer, and Miller (2004) and Moral-Benito (2012), we depart from a purely Bayesian approach to avoid having to specify priors for the parameters and keep results intuitive. Sala-i-Martin, Doppelhofer, and Miller (2004) suggest taking expectations of equation (3) with respect to β, and replacing the posterior mean of β. This approach is a Bayesian Averaging of Classical Estimates, or BACE. Here we follow Moral-Benito (2012, see also Dardanoni *et al.*, 2015), who extends the approach to maximum likelihood estimators—a Bayesian Averaging of Maximum Likelihood Estimates, or BAMLE:

(4) $$E(\beta|D) = \sum_{j=1}^{2^k} P(M_j|D)\, \hat{\beta}_{ML}^{j}$$

where $\hat{\beta}_{ML}^{j}$ is the maximum likelihood estimate of the coefficients in model M_j from a frequentist estimation. As equation 2 above shows, the ingredients needed to calculate the posterior model probability $P(M_j|D)$ are the prior model size $P(M_j)$ as well as (an approximation of) the integrated likelihood function $f(\mathrm{D}|M_j)$. We discuss both in turn in the following section.

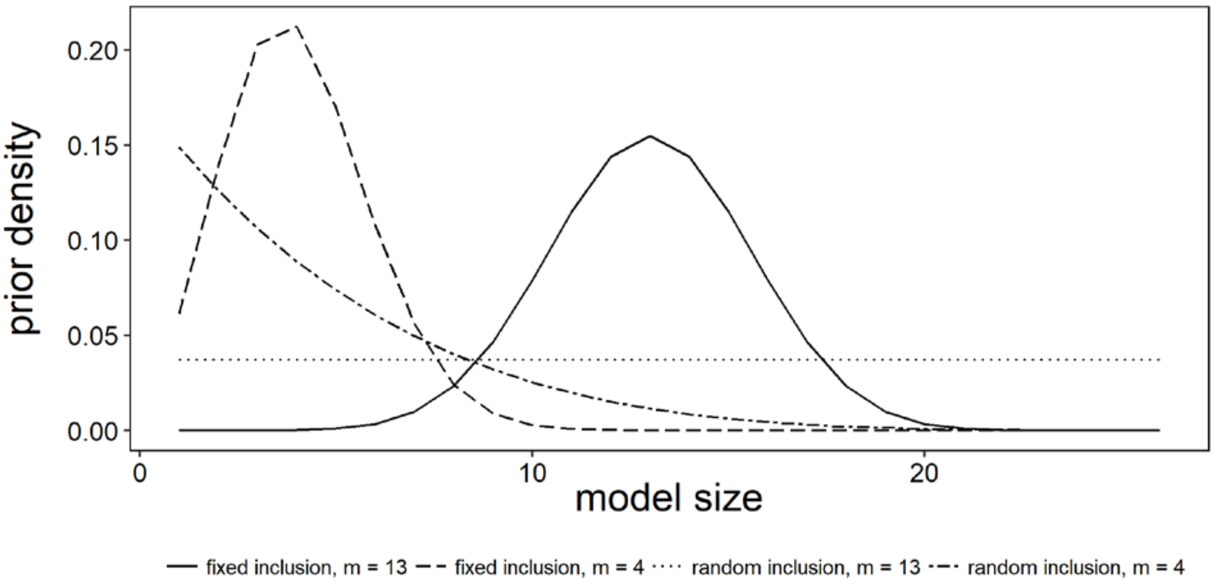
4.2 Binomial beta priors and the BIC approximation

One advantage of the BACE and BAMLE approaches is that only one prior, on either the inclusion probability of any given variable or the model size, is needed. Sala-i-Martin, Doppelhofer, and Miller (2004) employ only one prior on the inclusion probability θ of a specific variable which they reduce to $= \frac{m}{K}$, where m is the expected model size over which the researcher has a prior and K is again the total number of potentially relevant regressors. In a case of 26 potential variables and an expected model size of 4 variables, this would result in a prior inclusion probability of $\theta = 0.154$. No priors on the parameters are needed as the estimations are carried out in the frequentist world.

The priors used in the analysis should be minimal, unless we have very strong beliefs about them. The Sala-i-Martin, Doppelhofer, and Miller (2004) priors place relatively high prior density around the expected model size. Instead, Ley and Steel (2009) propose to further limit the information drawn from the prior by making it less informative, and draw the inclusion probability θ of each variable from a beta distribution $\theta \sim Be(a, b)$ with hyperparameters $a = 1$ and $b = \frac{K-m}{m}$ a function of the total number of estimators and the prior model size. With K potential variables, the first two moments of the resulting beta-binomial distribution of model size W are $E(W) = \frac{a}{a+b}K$ and $var(W) = \frac{ab(a+b+K)}{(a+b)^2(a+b+1)}K$ which are therefore a function of both m and K.

Again, the researcher only needs to set a prior on the mean model size which, together with the number of candidate regressors, determines the parameters and therefore the shape of the beta-binomial distribution. The difference with respect to the Sala-i-Martin, Doppelhofer, and Miller (2004) approach, however, is that this prior distribution has much less density around its mean, and is therefore less likely to dominate the results. In line with Moral-Benito (2012), we also run extensive robustness checks and confirm that our results are very robust to even extreme choices of m. To illustrate this point, Figure 4 plots the prior probabilities of the model size under this specification versus that of Sala-i-Martin, Doppelhofer, and Miller (2004), for our case of $K = 26$ regressors and priors on the model size of $m = 4$ as well as $m = \mathrm{K}/2$ (which would be the resulting prior from making no assumption on the model size). As can be seen, the binomial beta priors put less weight on models close to the prior, and are therefore less likely to drive the results. We run robustness checks for the prior specifications and use the Ley and Steel (2009) prior with random inclusion probability and $m = 4$ for the baseline specifications.

Figure 4. Binomial Beta priors



Note: Compares Sala-i-Martin, Doppelhofer, and Miller (2004) fixed inclusion probability at $\theta = \frac{m}{K}$ in the dashed and solid lines to the binomial beta priors suggested by Ley and Steel (2009) in the dash-dot and dotted lines for our sample size. The latter shows a strong reduction in prior confidence on the model space and therefore limits the effect of the prior inclusion probability θ.

Having established our choice of priors, the second ingredient needed to calculate the posterior model probabilities is the integrated likelihood function, or rather an approximation thereof. An approximation based on the Bayesian information criterion (BIC) is a natural choice for maximum likelihood estimates as the BIC is itself a likelihood based measure. Raftery (1995) provides the BIC approximation of the Bayes Factor, and Moral-Benito (2012) applies this approximation to BAMLE. Thus, $f\left(D\middle|M_j\right) \propto \exp(-\frac{1}{2}BIC_j)$ which, when applied to equation 2, gives the BIC approximation of the posterior model probability (for derivations, see Raftery, 1995):

(5) $$P\left(M_j\middle|\mathrm{D}\right) = \frac{P(M_j)\exp(-\frac{1}{2}BIC_j)}{\sum_{i=1}^{2^K} P(M_i)\exp(-\frac{1}{2}BIC_i)}$$

Therefore, the posterior probability $P\left(M_j\middle|\mathrm{D}\right)$ of model M_j is calculated using its prior probability and the BIC approximation over that same measure for all possible models. We can calculate this based on the priors $P\left(M_j\right)$ specified in the last section. These posterior model probabilities also allow us to calculate the posterior inclusion probability of each candidate regressor. It is simply the sum of the model probabilities of those models that include the candidate regressor:

(6) $$PiP(\beta^*) = \sum_{\beta^*=1}^{K^*} P(M_{\beta^*}|D)$$

where the star indicates that the specific variable in question is included in the model. This PiP can be compared to the naïve prior inclusion probability of a given variable pointed out above, in our case $\theta = m/K = 0.154$. However, this decision rule should not be applied mechanically (see Moral-Benito, 2012, for a discussion). We are also interested in the relative importance of the different variables. Therefore, we also discuss variables that only beat half the decision rule, or 0.077, when presenting our results.

4.3 Sampling

The large model space of 2^K possible variable combinations makes estimation of all models infeasible and motivates the use of a sampling algorithm that moves through a subset of the model space. Ideally, the sampling algorithm should limit the number of estimated ‘bad’ models that do not contribute much information. Following the exposition in Koop (2003), we implement a Markov Chain Monte Carlo Model Composition (MC[3]) routine as introduced by Madigan, York, and Allard (1995). This routine, which is commonly used in Bayesian econometrics, compares the current model M to a candidate model M^*. The candidate model M^* is drawn from a subset of the model space, namely the neighborhood consisting of model M, all possible models deleting one of the variables of M, and all models adding one variable to M. The candidate model M^* is accepted with probability $\alpha(M, M^*)$ equal to the posterior odds if these are smaller than unity:

(7) $$\alpha(M, M^*) = \min\left\{\frac{f(y|\, M^*)P(M^*)}{f(y|\, M)P(M)}\ ,\ \ 1\ \right\}$$

Since we do have a (weak) prior on the model space, the posterior odds do not reduce to the Bayes factor. The Markov Chain generated by the algorithm forms the model space used for our BAMLE exercise. A starting model is necessary to have a baseline M against which to compare M^*. To provide some intuition, say we let the algorithm start at the maximum model including all 26 variables used to investigate product market reforms. If our prior on the model size (4 variables) is reasonable, this model is very unlikely. In our case, the data supports this interpretation and the algorithm quickly converges to smaller model sizes as we show for all reform areas in Annex A1. Once the algorithm compares models in the area of the model space where the data dominates the results, it is very unlikely to return to these large models. However, by starting at 26 variables, we have forced the algorithm to calculate a number of models that it did not select based on the data. On order to eliminate the impact of these models and thereby reduce the importance of our starting value, we follow convention and discard the first 1% of all estimated models as a 'burn-in' sample that is not included in our results. Our results are very robust to different choices for the starting model and, as Annex A1 shows for all reform areas, this burn in sample is very generous. The rationale for employing MC^3 sampling here is to increase draws from regions of the model space where posterior probability is relatively high, thus decreasing the number of estimated models to be averaged. At the same time, such sampling eliminates the (infeasible) task of estimating the entire space of 2^K models[4]. We discuss robustness in other dimensions (such as prior choice) in the results section.

4.4 BAMLE in logit models explaining structural reforms

Having outlined our reasons for choosing a dichotomous reform variable when discussing structural reforms, we assume that the errors e of the underlying latent variable model $\mathrm{y}^* = \mathrm{x}\beta + e$ (with $y = I(y^* > 0)$ where y is the reform variable considered and $I()$ an indicator function) follow a standard logistic distribution:

(8) $$\mathrm{P}(\mathrm{y}_{\mathrm{it}} = 1|\mathrm{x}) = \mathrm{P}(\mathrm{y}_{\mathrm{it}}^* > 0|\mathrm{x}) = \Lambda(\mathrm{x}_{\mathrm{it}}\beta) = \frac{\exp(x\beta)}{1+\exp(x\beta)}$$

where $\Lambda()$ denotes the logit link function. Log odds ratios are not intuitive to interpret and difficult to compare across models based on different variable specifications. We therefore need to clarify what kind of estimates to compare across models, and which posterior model probability to employ.

[4] See Annex 1 on the properties of the algorithm employed.

Since estimates of the log odds ratio or the odds ratio reported by common statistical programs are not comparable across models beyond the signs of the coefficients, we employ average partial effects. Averaging across the distribution of all covariates included in the model considered leads to the following estimator which, unlike log odds ratios, is comparable across models (see for example Wooldridge, 2005a, 2005b):

(8) $$\hat{\beta}_j^{\mathrm{APE}} = \hat{\beta}_j N^{-1} \sum_{i=i}^{N} g(x_i \hat{\beta})$$

The second ingredient needed for the BAMLE procedure is the set of model weights, which are based on the BIC as outlined above. While the BIC is a likelihood-based measure and therefore extend naturally to logit models, its penalty function introduces room for value judgement. This is critical for BAMLE, as it is the only channel for the updating of our prior on model size (see Ley and Steel, 2009). The BIC calculates $BIC = -2(l) + \ln(N_j)\, k^*$ where l is the maximized likelihood function of model M_j, k^* is the number of parameters estimated and N_j the total number of observations. Two issues arise. First, the researcher has to decide whether to employ the number of countries as N_j, or instead assume independence across countries and use the total number of observations. We follow convention and use the latter. Second, in order to compare the BIC across different models with differing variables and parameters k^*, it is important that N_j doesn't change across models. This occurs for example because more data are available for the variables used in model M_j compared to model M_i or $N_j > N_i$. Therefore, we reduce our number observations to a balanced panel by keeping only those observations that are common across all K variables. The panel data we employ is therefore unbalanced across countries but balanced across variables.

Having outlined the BAMLE ingredients and the particularities of our logit application, we now apply these model averaging techniques to our logit model of the likelihood of major labor and product market reforms. In presenting the results, we show the PiP given in equation (6) but also, to facilitate comparisons with existing studies in the literature, the posterior means and standard deviations *conditional* on inclusion of the underlying model.

V. RESULTS

In this section, we present results from running the BAMLE method outlined above over our four reform indicators: reforms of PMR, EPL for regular contracts (regular EPL), EPL for temporary contracts (temporary EPL), and unemployment benefit generosity (replacement rate and duration). Section 5.1 discusses the main results. Section 5.2 checks for the robustness of our main results to different estimators and prior specifications, and addresses the multicollinearity issue hinted at in the introduction.

5.1 Main Results

Table 2 shows the detailed results of the BAMLE exercise for the four reforms and all regressors used in our analysis. For each of the four reforms, the first column shows the posterior mean of the average partial effect conditional on inclusion, the second its standard error conditional on inclusion and the third the posterior inclusion probability. We build our discussion primarily on the posterior inclusion probability, highlighting robustly correlated variables ($pip > 0.154$) in dark grey and still relatively robust variables ($0.154 > pip > 0.077$) in light grey (see Moral-Benito, 2012, for a discussion). The average partial effects are multiplied by 100 and are therefore interpretable as percentage point effects.

Table 2. BAMLE Results Over the Four Reform Areas

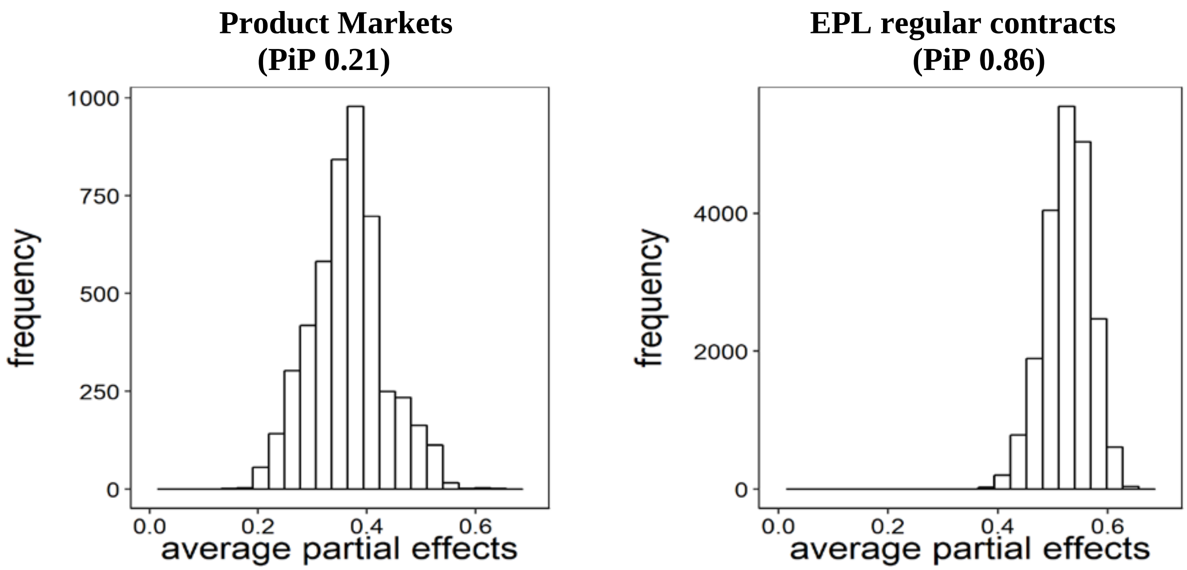
		Product Markets			EPL (temporary contracts)			EPL (regular contracts)			Unemployment Benefits		
		post.mean	**post.se**	**pip**	**post.mean**	**post.se**	**pip**	**post.mean**	**post.se**	**pip**	**post.mean**	**post.se**	**pip**
business conditions	gdp growth	-0.488	0.341	0.069	-0.298	0.382	0.045	-0.388	0.308	0.111	0.323	0.262	0.059
	unemployment	0.366	0.172	0.184	0.298	0.216	0.071	0.518	0.206	0.839	0.292	0.172	0.125
	deep recession	-1.575	1.707	0.037	2.516	0.922	0.924	1.108	0.834	0.086	-0.166	0.838	0.031
	crisis	2.499	3.304	0.038	1.305	2.735	0.032	1.446	2.404	0.040	-1.184	1.699	0.035
macroeconomic policies	exchange rate	0.271	0.208	0.059	0.101	0.231	0.035	0.081	0.223	0.032	0.236	0.172	0.070
	taylor rate	0.106	0.218	0.027	-0.526	0.327	0.135	-0.317	0.275	0.083	0.199	0.192	0.049
structural features of the domestic economy	openness	-2.439	1.533	0.093	-2.402	2.114	0.063	-1.734	2.173	0.042	1.181	0.760	0.045
	old age dep.	0.477	0.317	0.100	0.615	0.400	0.130	0.109	0.318	0.035	0.014	0.264	0.030
	gini (net)	-0.007	0.185	0.024	0.015	0.216	0.029	0.181	0.210	0.051	0.026	0.169	0.030
	gini (market)	0.155	0.154	0.045	0.253	0.247	0.066	0.148	0.188	0.048	0.154	0.152	0.045
	government debt	4.609	1.919	0.293	2.205	2.330	0.055	1.934	2.842	0.049	1.382	1.610	0.038
	regulation indicator	2.181	0.653	1	2.670	0.645	1	2.663	0.849	1	0.124	0.058	1
external factors	int. spillovers	0.988	0.277	0.998	0.902	0.389	0.696	0.058	0.509	0.032	0.050	0.363	0.031
	EU accession	9.651	6.749	0.134	3.305	6.260	0.033				3.709	6.622	0.037
	EU directives (pmr)	1.741	1.047	0.196									
political factors, including ideology, structural features of the political system, and conjunctural features (e.g. political cycles and crises)	exec r.l.c (cont)	-1.157	1.034	0.050	-0.158	0.947	0.023	-1.229	0.975	0.138	0.711	0.839	0.034
	exec left	-0.396	2.033	0.028	-0.607	1.763	0.028	-2.806	1.655	0.203	-1.074	1.817	0.032
	union density	-0.015	0.035	0.024	0.001	0.041	0.028	-0.072	0.052	0.149	0.024	0.029	0.041
	centralization gov.	1.750	2.915	0.031	2.534	3.206	0.048	-2.168	3.399	0.040	-2.270	2.808	0.040
	vote share gov.	-0.058	0.080	0.034	-0.074	0.046	0.063	-0.011	0.058	0.033	-0.006	0.059	0.029
	gov. controls all houses	-1.001	1.849	0.027	-0.551	1.936	0.028	-0.702	1.852	0.030	0.032	1.852	0.029
	months to election	-0.009	0.062	0.023	0.002	0.062	0.027	-0.017	0.063	0.045	0.010	0.049	0.029
	elections next year	1.099	1.839	0.026	-0.892	1.833	0.030	-3.147	1.555	0.224	0.346	1.477	0.030
	years left in office	-0.512	0.667	0.029	0.470	0.754	0.032	0.917	0.697	0.099	0.039	0.496	0.029
	years spent in office	0.252	0.242	0.044	-0.057	0.260	0.022	0.226	0.240	0.047	-0.207	0.244	0.044
momentum	domestic packaging	0.914	0.826	0.057	3.235	1.054	0.968	-0.473	0.763	0.041	0.168	0.679	0.030

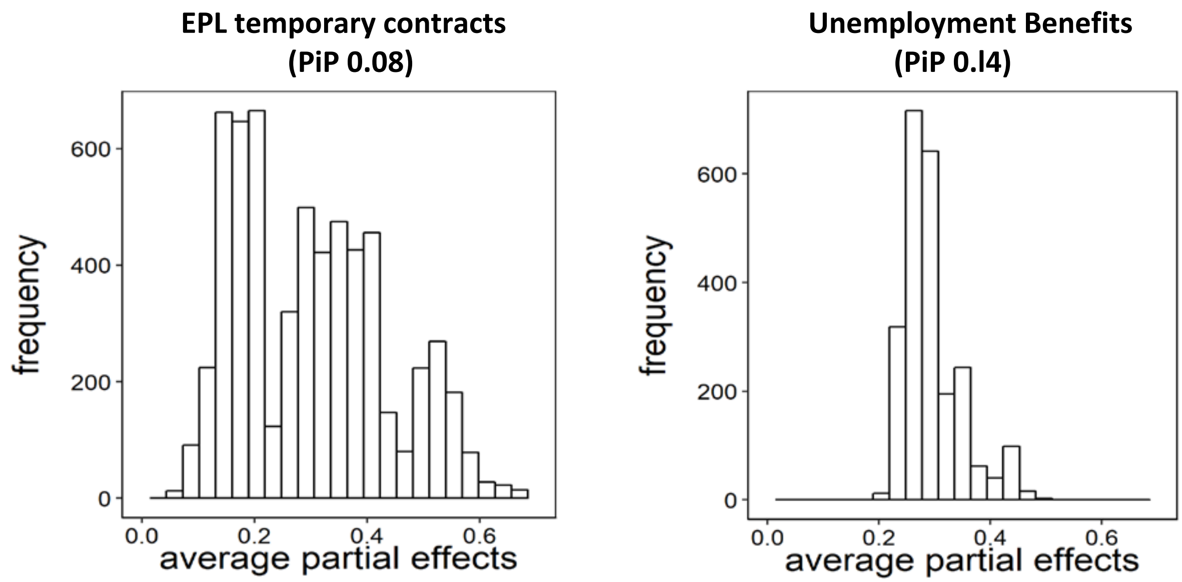
Note: Shows posterior means conditional on inclusion, standard errors conditional on inclusion, and posterior inclusion probabilities of the BAMLE exercise outlined in the methodological section. The table compares results of Logit models in four areas: product markets, employment protection legislation in temporary and regular contracts and unemployment benefits. These results can be interpreted as average partial effects for logit models. All results are based on binomial beta priors with m=4 as outlined in the methodological section.

n

The results of the analysis highlight several variables that appear robust across reforms. First and foremost, reforms across all areas tend to occur during period of high unemployment, low growth or recessions/crises, thereby supporting the crisis-induces-reform hypothesis (Table 2, Figure 5). The effects are economically significant. For example, an increase of 10 percentage points in unemployment (as seen in several European economies in the aftermath of the Great Recession) is associated with an increase in the probability to undertake a major regular EPL reform of about 5 percentage points—that is, about twice the average probability in the sample. Zooming in on the impact of unemployment, Figure 5 shows that the distributions of the average partial effect for all four reform areas are centered around positive means with practically no density in the negative space. The average partial effects converge to reasonably Gaussian distributions for product market reforms as well as regular EPL reforms, for which we found robust correlations. The distribution of average partial effects for temporary EPL is more spread out across the parameter space, but still on the right-hand side of zero; this reform area is the only one where unemployment is dominated by other variables measuring economic conditions—the deep recession variable.

Figure 5. Distribution of Average Partial Effects of Unemployment





Notes: The four panels show the frequency of average partial effects (in percentage points) of the first lag of the unemployment rate across the four reform areas taken from all unique models visited by the algorithm (1 million runs each). These effects should be interpreted conditional on inclusion with the posterior inclusion porbability given in parantheses for all four areas. The frequency on the vertical axis is made up of all unique models visited by the MC^3 algorithm that include the unemployment rate as a regressor. These results are based on one million iterations by the MC^3 algorithm for each reform area.

Second, average partial effects for the related regulation indicators are positive across the board. This is evidence that reforms are more likely to be undertaken when little action has been taken in the past, that is, when the scope for them is greater. For example, based on the posterior mean and standard error of the average partial effect of the regulatory stance (regulation indicator) conditional on inclusion, it appears that the likelihood of reform is higher when (lagged) product market regulation is more stringent.

Third, outside pressure increases the likelihood of reform in certain areas. Reforms in the areas of PMR and EPL for temporary contracts are more likely when other countries also undertake them. Formal pressure also matters: many product market reforms in EU countries have occurred during their accession process, and competition-relevant EU directives have also been an important factor behind deregulation.

In addition to these common factors, the analysis also points to a few important reform-specific determinants. Political factors play a significant role for reforms of EPL for regular workers, but not for other areas. In particular, the likelihood of regular EPL deregulation is lower under higher union density and left-leaning governments. This particular finding is consistent with theories that highlight the ability of entrenched interests to block reforms (e.g. Tommasi and Velasco, 1996). However, we do not find any evidence for a broader ideological bias—there is generally no robust difference between left- and right-of-center governments' propensity to undertake reform. The political business cycle also appears to have generally less importance than commonly assumed. Finally, another noticeable finding is that aging is associated with a higher likelihood of PMR and temporary EPL reforms, possibly because retirees reap gains but do not bear any of the potential costs—in terms of greater job insecurity, for example—from such reforms.

5.2 Robustness checks

Different model choices

To check the robustness of our results we apply the BAMLE methodology to include country and time fixed effects (Moral-Benito, 2012), and also consider a linear probability model. We refrained from including fixed effects in the main specification for two reasons. First, a model averaging exercise seeks to make inferences about the relationship between potentially relevant variables without imposing much structure. Second, in a logit set-up the inclusion of country (time) fixed effects would automatically drop the country (time) for which a given reform has never occurred.

The BAMLE results for a linear probability model as well as with country and time fixed effects are reported for regular EPL reforms in Table 3; results for the other reform areas yield similar conclusions and can be found in Annex 2 (Tables A2.1-A2.3). The first three columns show our baseline (pooled logit) results. Columns 4-6 repeat this exercise for a linear probability model, with the posterior mean then being directly derived from the estimates as average partial effects are no longer necessary. This has little impact on our results; in particular, the unemployment rate is still the most robust correlate of regular EPL reforms.

Table 3. BAMLE Results for Different Estimators (EPL regular)

		Logit pooled		Linear Prob. Model pooled		Logit time and year demeaned		Linear Prob. Model time and year demeaned	
		sign	pip	sign	pip	sign	pip	sign	pip
business conditions	gdp growth	-	0.111	-	0.117	-	0.175	-	0.437
	unemployment	+	0.839	+	0.942	+	0.083	+	0.086
	deep recession	+	0.086	+	0.159	+	0.133	+	0.139
	crisis	+	0.040	+	0.017	+	0.023	+	0.028
macroeconomic policies	exchange rate	+	0.032	+	0.050	+	0.016	+	0.029
	taylor rate	-	0.083	-	0.094	+	0.023	+	0.017
structural features of the domestic economy	openness	-	0.042	-	0.014	-	0.163	-	0.015
	old age dep.	+	0.035	-	0.011	-	0.011	-	0.015
	gini (net)	+	0.051	+	0.098	-	0.029	-	0.020
	gini (market)	+	0.048	+	0.025	+	0.012	+	0.035
	government debt	+	0.049	+	0.013	+	0.037	+	0.044
	regulation indicator	+	1	+	1	+	1	+	1
external factors	int. spillovers	+	0.032	+	0.032				
political factors, including ideology, structural features of the political system, and conjunctural features (e.g. political cycles and crises)	exec r.l.c (cont)	-	0.138	-	0.098	-	0.108	-	0.103
	exec left	-	0.203	-	0.080	-	0.120	-	0.209
	union density	-	0.149	-	0.306	+	0.020	-	0.022
	centralization gov.	-	0.040	+	0.016	-	0.060	-	0.111
	vote share gov.	-	0.033	-	0.021	+	0.028	+	0.017
	gov. controls all houses	-	0.030	-	0.016	-	0.020	-	0.019
	months to election	-	0.045	-	0.015	-	0.022	-	0.023
	elections next year	-	0.224	-	0.097	-	0.131	-	0.115
	years left in office	+	0.099	+	0.037	+	0.075	+	0.079
	years spent in office	+	0.047	+	0.017	+	0.022	+	0.035
momentum	domestic packaging	-	0.041	-	0.024	-	0.022	-	0.021

Note: Shows posterior means conditional on inclusion, standard errors conditional on inclusion, as well as posterior inclusion probabilities of the BAMLE exercise discussed in the methodological section over different estimators. The first three result columns repeat the baseline results based on Logit models. Columns 4-6 show the same exercise based on a linear probability model. Columns 7-9 show results of Logit models run on time and individually demeaned data.

Columns 7-9 then show results of logit models demeaned along the country and time dimensions. Several features stand out. First, we cannot include international spillovers anymore since, apart from a country's own reforms, the demeaned spillover variables are identical across countries. Next, we see that as we neglect the level information of the unemployment rate, GDP growth shows a higher pip and a still negative effect. This change in ordering suggests that two aspects of economic conditions drive regular EPL reforms, namely structurally poor labor market performance and weak business cycle conditions; countries with higher unemployment are more likely to implement regular EPL reforms and, once this level effect is controlled for, the business cycle—as captured by GDP growth—also matters. The last three columns of Table 3 confirm these results based on estimations of linear probability models with country and time fixed effects. All in all, we see that except for the unemployment rate—for clear reasons—the results are surprisingly consistent across these changes in the estimators.

Addressing multicollinearity

Another potential issue with our analysis, as well as most papers in the field, is that a number of the variables we consider are strongly correlated, raising a collinearity problem. For example, both unemployment and interest rate spreads will increase during recessions. In this section, instead of choosing amongst them on an ad-hoc basis and excluding the others, as is common in the growth literature, we provide an alteration of the BAMLE methodology where we restrict the model space to never include such variables together in the same model while still using results for all variables. The following restriction sets are introduced, meaning that at most one variable of each restriction set (bullet point) is included in the same model:

- GDP growth, deep recession, unemployment, crisis
- The two Gini coefficients (net and market)
- Continuous ideology variable and dummy for leftwing executive
- Government vote share and government party majority in all houses
- Number of months to elections, close elections, years left in office, years executive has already been in office

We restrict the algorithm moving through the model space in the following way. First, we start with a model that uses the maximum number of variables, before discarding all but one randomly drawn variable from each of the five restriction sets. This starting model will, for example, not include deep recession and crisis variables together. The updating process then runs as before for smaller and identical models. However, when the algorithm moves to a larger model, we need to ensure that this larger model does not include more than one variable from each restriction set. To achieve this, we restrict the set of variables the algorithm draws from to the set of variables that does not violate the restriction sets.

Table 4. Addressing Multicollinearity with Exclusion Sets

		Product Markets		EPL (temporary contracts)		EPL (regular contracts)		Unemployment Benefits	
		pip baseline	pip restricted	pip baseline	pip restricted	pip baseline	pip restricted	pip baseline	pip restricted
business conditions	gdp growth	0.069	0.047	0.045	0.009	0.111	0.016	0.059	0.031
	unemployment	0.184	0.141	0.071	0.011	0.839	0.794	0.125	0.093
	deep recession	0.037	0.008	0.924	0.913	0.086	0.024	0.031	0.015
	crisis	0.038	0.023	0.032	0.001	0.040	0.004	0.035	0.020
macroeconomic policies	exchange rate	0.059	0.042	0.035	0.015	0.032	0.020	0.070	0.042
	taylor rate	0.027	0.018	0.135	0.102	0.083	0.056	0.049	0.032
structural features of the domestic economy	openness	0.093	0.070	0.063	0.032	0.042	0.021	0.045	0.028
	old age dep.	0.100	0.065	0.130	0.103	0.035	0.020	0.030	0.016
	gini (net)	0.024	0.014	0.029	0.012	0.051	0.029	0.030	0.016
	gini (market)	0.045	0.026	0.066	0.040	0.048	0.039	0.045	0.026
	government debt	0.293	0.239	0.055	0.049	0.049	0.033	0.038	0.020
	regulation indicator	1	1	1	1	1	1	1	1
external factors	int. spillovers	0.998	1.000	0.696	0.655	0.032	0.016	0.031	0.017
	EU accession	0.134	0.094	0.033	0.016			0.037	0.026
	EU directives (pmr)	0.196	0.130						
political factors, including ideology, structural features of the political system, and conjunctural features (e.g. political cycles and crises)	exec r.l.c (cont)	0.050	0.037	0.023	0.014	0.138	0.127	0.034	0.019
	exec left	0.028	0.021	0.028	0.016	0.203	0.171	0.032	0.021
	union density	0.024	0.019	0.028	0.026	0.149	0.126	0.041	0.028
	centralization gov.	0.031	0.021	0.048	0.031	0.040	0.022	0.040	0.031
	vote share gov.	0.034	0.027	0.063	0.037	0.033	0.020	0.029	0.016
	gov. controls all houses	0.027	0.015	0.028	0.021	0.030	0.022	0.029	0.018
	months to election	0.023	0.017	0.027	0.008	0.045	0.023	0.029	0.014
	elections next year	0.026	0.015	0.030	0.012	0.224	0.168	0.030	0.018
	years left in office	0.029	0.026	0.032	0.030	0.099	0.070	0.029	0.019
	years spent in office	0.044	0.030	0.022	0.021	0.047	0.021	0.044	0.029
momentum	domestic packaging	0.057	0.040	0.968	0.969	0.041	0.020	0.030	0.016

Note: This table shows posterior inclusion probabilities of the BAMLE exercise under two different specifications for the four reform areas. For each reform area, the first column (‘baseline’) shows the posterior inclusion probability in our baseline analysis. The second column (‘restricted’) shows the results for a BAMLE estimation with the restriction sets outlined in the main text. The dark shaded inclusion probabilities are larger than the prior inclusion probabilities (> 0.154) while the light shaded areas indicate variables that meet half of this criterion (> 0.077).

Table 4 shows the results from this analysis, comparing for each reform area the baseline results with those from the restricted version. Key findings remain largely unaffected by this adjustment, which shows that the BAMLE methodology is highly robust, even to multicollinearity issues.

Changing prior specifications

As a final sensitivity analysis, we test the robustness of our results to the choice of the prior mean model size m. We focus here on regular EPL reforms, but results for other reform areas (shown in Annex 3, Table A3.1) deliver yield the same message. Table 5 shows the ten variables with the highest posterior inclusion probability (pip) of BAMLE exercises based on three alternative choices for m. Despite the drastic variation in the prior, all three specifications produce comparable results.

Table 5. Robustness of Results to Extreme Prior Choices

Panel 1, prior: $m = 1$

	pip
regulation indicator	1
unemployment	0.988
elections next year	0.578
exec left	0.518
union density	0.254
exec r.l.c (cont)	0.207
gdp growth	0.193
taylor rate	0.156
years left in office	0.148
years spent in office	0.115

Panel 2, prior: $m = K/2$

	pip
regulation indicator	1
unemployment	0.930
elections next year	0.301
exec left	0.278
union density	0.194
exec r.l.c (cont)	0.151
gdp growth	0.102
taylor rate	0.094
deep recession	0.090
years left in office	0.079

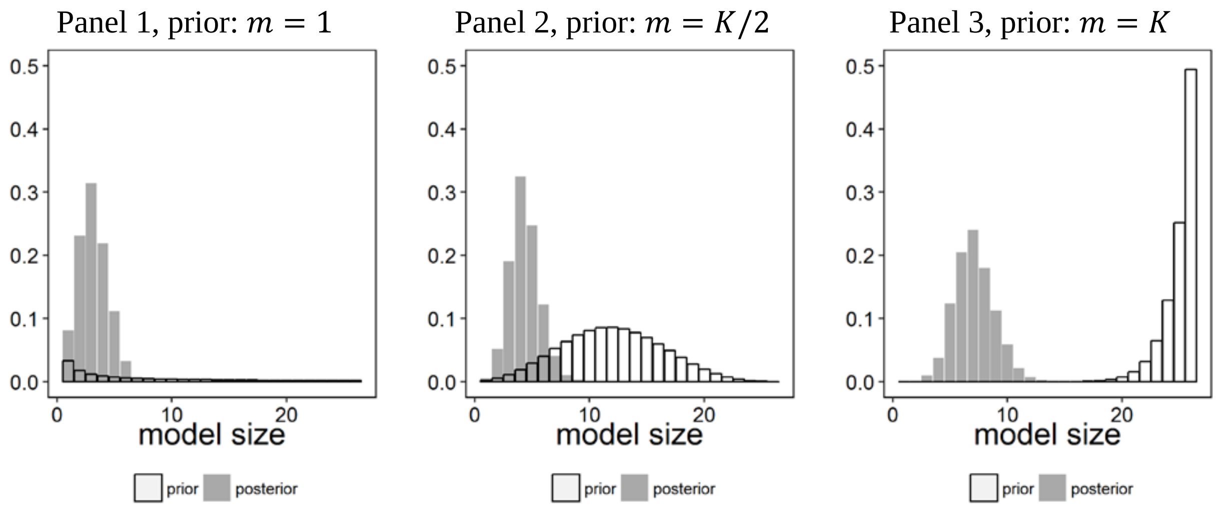
Panel 3, prior: $m = K$

	pip
regulation indicator	1
unemployment	0.988
elections next year	0.578
exec left	0.518
union density	0.254
exec r.l.c (cont)	0.207
gdp growth	0.193
taylor rate	0.156
years left in office	0.148
years spent in office	0.115

Notes: Each panel shows the 10 variables with the highest posterior inclusion probability (pip) as well as their pip for different choices of the hyperparameter m. In the first panel, we choose the lower extreme (prior: $m = 1$). In the second panel, we assume that a model with a model size of half of the available variables is the most likely model (prior: $m = K/2$). In the third panel, we consider the maximum number of variables as the most likely model (prior: m=K). All three panels show the same ordering for the first five variables with some minor shifts (by one position) for the next five variables. This shows the strong robustness of our results to choices of the hyperparameter m. These results are based on calculations analyzing reforms of employment protection legislation in regular contracts but results for other reforms show the same picture.

We also check the behavior of our algorithm with respect to prior choices regarding model size. The three panels of figure 6 show extreme prior choices. The grey bars show the distribution of the posterior model size, while the white bars show that of the priors. Although the mode of the prior moves from one extreme to the other from panel 1 through panel 3, the posterior consistently indicates relatively moderate model sizes of around six to seven variables. Figure 6 thus shows that the posterior model size is dominated by the data, not by our assumed prior. This important finding for the robustness of our results is in line with those in Ley and Steel (2009).

Figure 6. Robustness of Posterior Model Size to Extreme Prior Choices



Notes: The barcharts show prior (empty framed bars) and posterior (grey bars) model sizes for three different choices of hyperparameter m. In the first panel, we choose the lower extreme (prior: $m = 1$). In the second panel, we assume that a model with a model size of half of the available variables is the most likely model (Prior: $m = K/2$). In the third panel, we consider the maximum number of variables as the most likely model. In all of these cases, the data dominates the results as discussed in the main text and visible in the grey posterior bars. These results are based on calculations analyzing reforms of employment protection legislation in regular contracts. Results for the other three reform areas look similar and are shown in Annex 2.

VI. CONCLUSION

This paper provides the first attempt in the literature on the drivers of structural reforms to address model uncertainty head-on, while also improving on the identification of reforms through a new “narrative” approach. Both contributions aim to enhance our knowledge of reform drivers against the background of very limited consensus thus far. Potential determinants identified in the literature have included crisis episodes, economic conditions, trade openness, policy diffusion, reform momentum, fiscal space, accommodative monetary policy, political capital, political ideology, democracy, institutional structure, government or parliamentary instability and many more. In addition, major reforms have often been identified indirectly from available policy indicators using largely *ad hoc* criteria, rather than directly based on information about the timing and nature of actual regulatory and legislative changes. Our reform identification approach, which relies on a new “narrative” database of labor and product market reforms in advanced countries, helps to address the latter source of uncertainty. Using this database, we then tailor a Bayesian averaging of maximum likelihood estimators (BAMLE) approach to binary logit models, which allows us to test a large list of potential drivers simultaneously.

Our results confirm the crisis-induces-reform hypothesis for all four reform areas we analyze—product market deregulation in network industries, reforms that ease EPL for either regular or temporary contracts, and reductions in the generosity of unemployment benefits. The relative importance of indicators of weak economic conditions differs somewhat across reform areas; product market deregulation and unemployment benefit cuts are mostly associated with periods of high unemployment, temporary EPL reform is mostly passed during deep recessions, while regular EPL reforms are enacted under all variants of weak economic conditions. Other robust reform drivers include the scope for deregulation and international pressure (PMR). Finally, a few political factors are robustly associated with reforms of EPL for regular contracts. We also establish an interesting list of non-robust results. In particular, we cannot confirm the it-takes-a-Nixon-to-go-to-China hypothesis: with the exception of regular EPL reforms, the political orientation of the executive does not emerge as a robust driver of structural reform. We also find little support for a political business cycle argument: the timing of elections seems to have at most a minor effect on the likelihood of reform.

With this paper, we hope to contribute to a better understanding of the drivers of structural reforms, discriminating between the multitude of hypotheses that have been advanced in the literature. Our results open interesting venues of future research. In part for data availability reasons, our focus here has been on labor and product market reforms. However, there is scope for similar analysis of the drivers of reforms in other areas, and for identifying the latter using the type of narrative approach followed in the present paper. In particular, given its prominence and controversy in the literature, it would be useful to confirm the strength of the crisis-induces-reform hypothesis across a broader array of reforms. Also, while the BAMLE approach is a statistical exercise that provides little to no information on the channels that underpin the results, it helps identify fruitful areas for further theoretical and empirical research. Our results highlight several of these.

VII. REFERENCES

Abiad, Abdul and Ashoka Mody, 2005. Financial Reform: What shakes it? What shapes it? *American Economic Review* 95(1): 66-88.

Acemoglu, Daron, Simon Johnson, and James Robinson, 2006. Institutions as the fundamental cause of long-run growth. In: Philippe Aghion and Steven N. Durlauf (Eds.), *Handbook of Economic Growth*. Elsevier, Amsterdam: 385–472.

Agnello, Luca, Vitor Castro, João T. Jalles, and Ricardo M. Sousa, 2015. What determines the likelihood of structural reforms? *European Journal of Political Economy* 37(1): 129-145.

Alesina, Alberto, Silvia Ardagna, and Francesco Trebbi, 2006. *Who adjusts and when? On the political economy of reforms*. NBER Working Paper No. 12049.

Alesina, Alberto and Allan Drazen, 1991. Why are stabilizations delayed? *American Economic Review* 81(5): 1170-1188.

Bates, Robert H. and Anne O. Krueger, 1993. Generalizations arising from the country studies. In: Robert H. Bates and Anne O. Krueger (Eds.), *Political and Economic Interactions in Economic Policy Reform: Evidence from Eight Countries*. Basil Blackwell, Oxford.

Bean, Charles, 1998. The interaction of aggregate demand policies and labor market reform. *Swedish Economic Policy Review* 5(2): 353-382.

Belloc, Filippo, and Antonio Nicita, 2011. The political determinants of liberalization: do ideological cleavages still matter? *International Review of Economics* 58: 121-145.

Blanchard, Olivier, and Francesco Giavazzi, 2003. Macroeconomic effects of regulation and deregulation in goods and labor markets. *Quarterly Journal of Economics* 118(3): 879–907.

Bonfiglioli, Allessandra, and Gino Gancia, 2013. Uncertainty, electoral incentives and political myopia. *Economic Journal* 123(568): 373-400.

Bouis, Romain, Romain Duval, and Johannes Eugster, 2016. *Product market deregulation and growth: New country-industry-level evidence*. IMF Working Paper No. WP/16/114.

Bortolotti, B., Marcella Fantini, and Domenico Siniscalco, 2003. Privatization around the world: evidence from panel data. *Journal of Public Economics* 88: 305–332.

Bortolotti, Bernardo, and Paolo Pinotti, 2008. Delayed privatization. *Public Choice* 136(3): 331-351.

Bruno, Michael and William Easterly, 1996. Inflation's children: Tales of crises that beget reforms. *American Economic Review Papers and Proceedings* 86(2): 213–217.

Cacciatore, Matteo, Romain Duval, Giuseppe Fiori, and Fabio Ghironi, 2016. Short-term pain for long-term gain: Market deregulation and monetary policy in small open economies. *Journal of International Money and Finance* 68: 358–385.

Campos, Nauro F., Cheng Hsiao, and Jeffrey B. Nugent, 2010. Crises, what crises? New evidence on the relative roles of political and economic crises in begetting reforms. *Journal of Development Studies* 46(10): 1670-1691.

Campos, Nauro F., and Roman Horvath, 2012a. Reform redux: Measurement, determinants and growth implications. *European Journal of Political Economy* 28(2): 227-237.

Campos, Nauro F., and Roman Horvath, 2012b. On the reversibility of structural reforms. *Economics Letters* 117(1): 217-219.

Cukierman, Alex, and Mariano Tommasi, 1998. When does it take a Nixon to go to China? *American Economic Review* 88(1): 180-197.

Dardanoni, Valentino, Guiseppe De Luca, Salvatore Modica, and Franco Peracchi, 2015. Model averaging estimation of generalized linear models with imputed covariates. *Journal of Econometrics* 184(2): 452-463.

Draghi, Mario, 2016. The productivity challenge for Europe. Lecture at the Deusto Business school, 30.11.2016. Available at: https://www.ecb.europa.eu/press/key/date/2016/html/sp161130_1.en.html

Drazen, Allan, 2000. *Political Economy in Macroeconomics*. Princeton University Press: Princeton, NJ.

Drazen, Allan, and William Easterly, 2001. Do crises induce reform? Simple empirical tests of conventional wisdom. *Economics and Politics* 13(2): 129-157.

Drazen, Allan and Vittorio Grilli, 1993. The benefit of crises for economic reforms. *American Economic Review* 83(3): 598–607.

Durlauf, Steven N., Chau Fu, and Salvador Navarro, 2012. Assumptions matter: Model uncertainty and the deterrent effect of capital punishment. *American Economic Review: Papers and Proceedings* 102(3): 487-492.

Duval, Romain, 2008. Is there a role for macroeconomic policy in fostering structural reforms? Panel evidence from OECD countries over the past two decades. *European Journal of Political Economy* 24(2): 491-502.

Duval, Romain, and Jörgen Elmeskov, 2006. *The effects of EMU on structural reforms in labor and product markets*. ECB working paper No. 596.

Duval, Romain and Davide Furceri, 2018. The effects of labor and product market reforms: The role of macroeconomic conditions and policies. *IMF Economic Review,* Feb 2018.

Duval, Romain, Davide Furceri, Bingjie Hu, Joao Jalles, and Huy Nguyen, 2018. *A narrative database of major labor and product market reforms in advanced economies*. IMF Working Paper No. WP/18/19.

Eichengreen, Barry,Charles Wyplosz, Charles Bean, and Stefan Gerlach, 1998. The stability pact: More than a minor nuisance? *Economic Policy* 13(26): 65–113.

Elhorst, Paul, Eelco Zandberg, and Jakob De Haan, 2013. The impact of interaction effects among neighbouring countries on financial liberalization and reform: A dynamic spatial panel data approach. *Spatial Economic Analysis* 8(3): 293-313.

Fernandez, Raquel and Dani Rodrik, 1991. Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review* 81(5): 1146–1155.

Fidrmuc, Jan and Elira Karaja, 2013. Uncertainty, informational spillovers and policy reform: A gravity model approach. *European Journal of Political Economy* 32: 182–192.

Galasso, Vincenzo, 2006. *The Political Future of Social Security in an Ageing Society*. The MIT Press: Cambridge, Massachussets.

Galasso, Vincenzo, 2014. The role of political partisanship during economic crises. *Public Choice* 158(1): 143-165.

Gassebner, Martin, Noel Gaston, and Michael Lamla, 2011. The inverse domino effect: are economic reforms contagious? *International Economic Review* 52(1): 183–200.

Giuliano, Paola, Prachi Mishra, and Antonio Spilimbergo, 2013. Democracy and Reforms: Evidence from a New Dataset. *American Economic Journal: Macroeconomics* 5(4): 179-204.

Golinelli, Roberto and Riccardo Rovelli, 2013. Did growth and reforms increase citizens' support for the transition? *European Journal of Political Economy* 30: 112–137.

Grindle, Merilee and John W. Thomas, 1991. *Public Choices and Policy Change: The Political Economy of Reform in Developing Countries*. Johns Hopkins University Press, Baltimore and London.

Gupta, Sanjeev, Estelle X. Liu, and Carlos Mulas-Granados, 2016. Now or later? The political economy of public investment in democracies. *European Journal of Political Economy* 45: 101-114.

Haggard, Stefan and Robert R. Kaufman, 1992. *The Politics of Economic Adjustment: International Constraints, Distributive Conflicts, and the State*. Princeton University Press, Princeton.

Haggard, Stefan and Steven B. Webb, 1994. *Voting for Reform: Democracy, Political Liberalization and Economic Adjustment*. Oxford University Press: New York.

Hoj, Jens, Vincenzo Galasso, Giuseppe Nicoletti and Thai-Thanh Dang, 2007. An empirical investigation of political economy factors behind structural reforms in OECD countries. *OECD Economics Studies* 2006(1): 87-136.

Hjort, Nils L. and Gerda Claeskens, 2003. Frequentist model average estimators. *Journal of the American Statistical Association* 98(464): 879-899.

Hjort, Nils L. and Gerda Claeskens, 2008. *Model selection and model averaging*. Cambridge University Press, Cambridge.

Ilzetzki, Ethan, Carmen M. Reinhart, and Kenneth S. Rogoff, 2017. *Exchange arrangements entering the 21st century: Which anchor will hold?* NBER Working Paper No. 23134.

International Monetary Fund (IMF), 2004. Fostering structural reforms in industrial countries. *World Economic Outlook April 2004, Chapter III*. IMF, Washington DC.

Koop, Gary, 2003. *Bayesian Econometrics*. John Wiley & Sons Ltd., Chichester.

Krueger, Anne O., 1993. *Political Economy of Policy Reform in Developing Countries*. MIT Press, Cambridge, MA.

Leamer, Edward E., 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. Wiley, New York.

Leamer, Edward E., 1985. Sensitivity analysis would help. *American Economic Review* 75(3): 308-313.

Levine, Ross and David Renelt, 1992. A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82(4): 942-963.

Ley, Eduardo, and Mark F. J. Steel, 2009. On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics* 24(4): 651-674.

Li, Meng and David B. Dunson, 2017. *A framework for probabilistic inferences from imperfect models*. Unpublished Working Paper.

Lora, Eduardo A. and Mauricio Olivera, 2004. What makes reforms likely: Political economy determinants of reforms in Latin America. *Journal of Applied Economics* 7(1): 99–135.

Madigan, David, Jeremy York, and Denis Allard, 1995. Bayesian graphical models for discrete data. *International Statistical Review* 63(2): 215-232.

Martinelli, César, and Mariano Tommasi, 1995. *Economic reforms and political constraints: On the time inconsistency of gradual sequencing*. University of California Working Paper No. 736.

Meseguer, Covadonga, 2006. Learning and economic policy choices. *European Journal of Political Economy* 22(1): 156–178.

Mian, Atif, Amir Sufi, and Francesco Trebbi, 2014. Resolving debt overhang: Political constraints in the aftermath of financial crises. *American Economic Journal: Macroeconomics* 6(2): 1-28.

Moral-Benito, Enrique, 2012. Determinants of economic growth: A Bayesian panel data approach. *Review of Economics and Statistics* 94(2): 566-579.

Moral-Benito, Enrique, 2015. Model averaging in economics: An overview. *Journal of Economic Surveys* 29(1): 46-75.

Nelson, Joan M., 1990. *Economic Crisis and Policy Choice: The Politics of Adjustment in the Third World*. Princeton University Press, Princeton.

Olson, Mancur, 1965. *The Logic of Collective Action*. Harvard University Press, Cambridge, MA.

Persson, Torsten and Guido Tabellini, 1994. Is inequality harmful for growth? *American Economic Review* 84(3): 600-621.

Persson, Torsten and Guido Tabellini, 2000. *Political Economics: Explaining Economic Policy*. MIT Press, Cambridge, MA.

Pitlik, Hans and Steffen Wirth, 2003. Do crises promote the extent of economic liberalization?: An empirical test. *European Journal of Political Economy*, 19(3): 565-581.

Potrafke, Niklas, 2009. Did globalization restrict partisan politics? An empirical evaluation of social expenditures in a panel of OECD countries. *Public Choice* 140(1): 105-124.

Raftery, Adrian E., 1995. Bayesian model selection in social research. *Sociological Methodology* 25: 111-163.

Roberts, Barbara M. and Muhammad A. Saeed, 2012. Privatizations around the world: Economic or political determinants? *Economics and Politics* 24(1): 47-71.

Rodrik, Dani, 1996. Understanding economic policy reform. *Journal of Economic Literature* 34(1): 9–41.

Romer, Christina D., and David H. Romer, 2010. The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review* 100(3): 763–801.

Romer, Christina D. and David H. Romer, 2017. New evidence on the aftermath of financial crises in advanced countries. *American Economic Review* 107(10): 3072–3118.

Saint-Paul, Gilles, 1997. The Rise and Persistence of Rigidities. *American Economic Review Papers and Proceedings* 87(2): 290-294.

Saint-Paul, Gilles, 2000. *The Political Economy of Labor Market Institutions*. Oxford University Press: Oxford.

Sala-i-Martin, Xavier X., 1997. I just ran two million regressions. *American Economic Review* 87(2): 178-183.

Sala-i-Martin, Xavier X., Gernot Doppelhofer, and Ronald I. Miller, 2004. Determinants of long-term growth: A Bayesian Averaging of Classical Estimates (BACE) approach. *American Economic Review* 94(4): 813-835.

Smets, Lodewijk, Stephen Knack, and Nadja Molenaers, 2012. Political ideology, quality at entry and the success of economic reform programs. *Review of International Organizations* 8(4): 447-476.

Tommasi, Mariano and André Velasco, 1996. Where are we in the political economy of reform? *Journal of Policy Reform* 1(2): 187-238.

Valencia, Luc, and Fabián Laeven, 2012. *Systemic banking crises database: An update*. IMF Working Paper No. WP/12/163.

Volden, Craig, Michael M. Ting, and Daniel P. Carpenter, 2008. A formal model of learning and policy diffusion. *American Political Science Review* 102(3): 319–332.

Wiese, Rasmus, 2014. What triggers reforms in OECD countries? Improved reform measurement and evidence from the healthcare sector. *European Journal of Political Economy* 34: 332-352.

Williamson, John, 1990. *Latin American Adjustment: How Much Has Happened?* Institute for International Economics, Washington, DC.

Williamson, John and Stefan Haggard, 1994. The political conditions for economic reform. In: John Williamson (Ed.), *The Political Economy of Policy Reform*. Institute for International Economics, Washington, DC, 525–596.

Wooldridge, Jeffrey M., 2005a. Unobserved heterogeneity and estimation of average partial effects. In: Donald W. K. Andrews and James H. Stock (Eds.), *Identification and Inference for Econometric Models. Essays in Honor of Thomas Rothenberg*. Cambridge University Press, New York: 27-55.

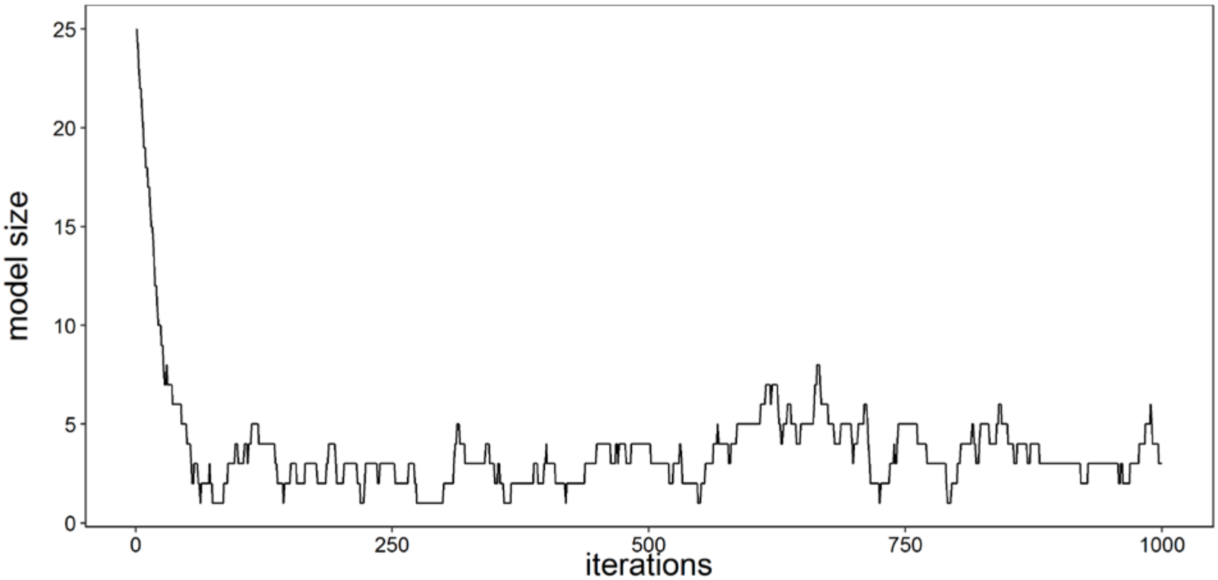
Wooldridge, Jeffrey M., 2005b. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20(1): 39-54.

Annex 1. Properties of the Sampling Algorithm

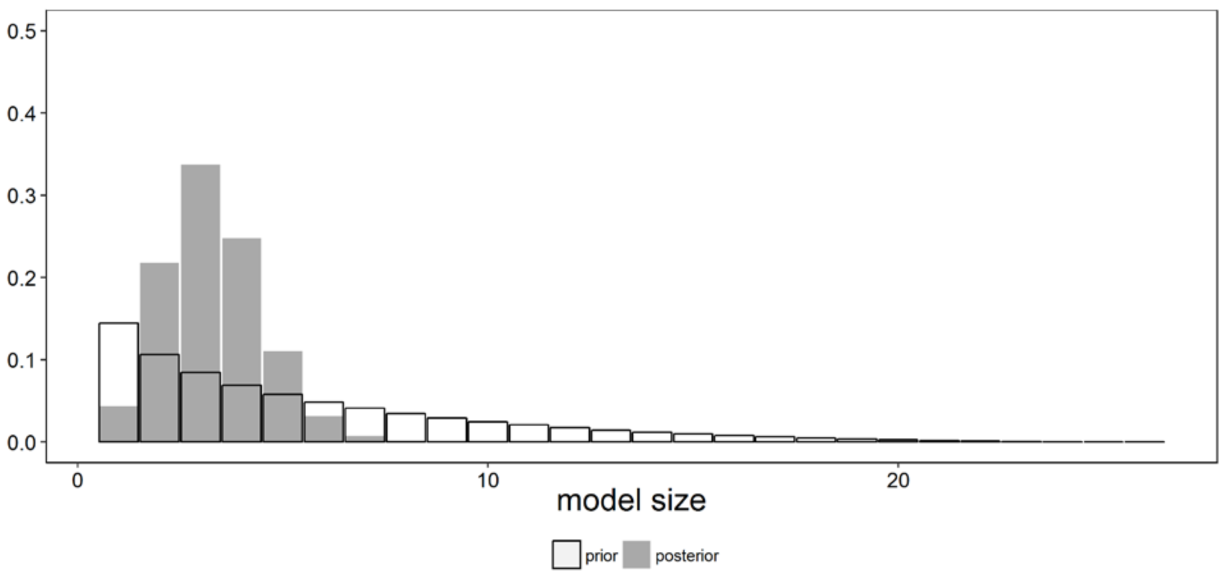
In order to verify the properties of the algorithm employed, Figure A1.1 shows its behavior when used to analyze reforms of EPL for regular contracts. Panel 1 shows the first 1000 iterations of the algorithm when we force it to start at the maximum model size. Within 200 iterations, the algorithm converges towards reasonable model sizes within a range between two and seven variables per model. This suggests a required burn in of only around 200 iterations. Once this region of the model space is reached, the algorithm stays within these bounds for the most part. However, the algorithm does not converge to a single model size that would be needed to select the one best model to explain structural reforms. Model uncertainty is too high for such a conclusion, which again points to the importance of employing a model averaging estimator to investigate correlations with structural reforms. Panel 2 compares the prior and the posterior of the model size. As Ley and Steel (2009) note, the prior needs to be dominated by the data to show that any sensible model exists: the mode of the posterior should therefore be larger than the smallest model if the data is informative. This is the case here. The posterior distribution of the model size looks reasonably Gaussian with a mean of 3.4 and, for regular EPL reforms as used in Figure 3, it actually updates our belief of the mean model size to a smaller number than the proposed four variables. Results on the other three reform areas show similar results.

Figure A1.1. Behavior of the MC³ algorithm

Panel 1: Convergence of the MC³ algorithm: EPL regular contracts

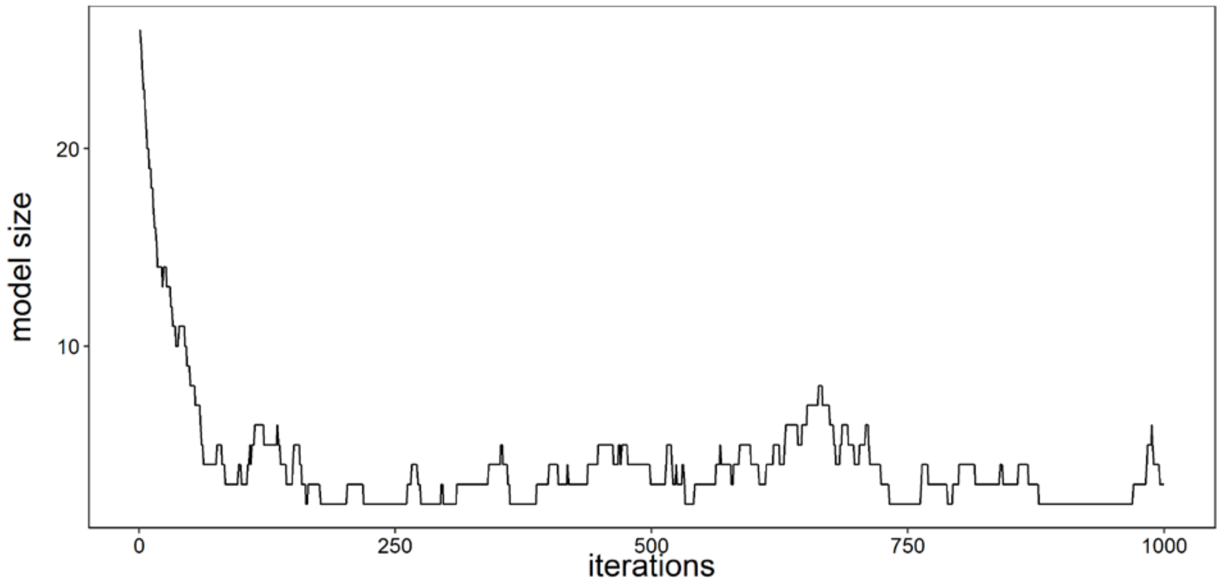


Panel 2. Prior and Posterior distribution over model size: EPL regular contracts

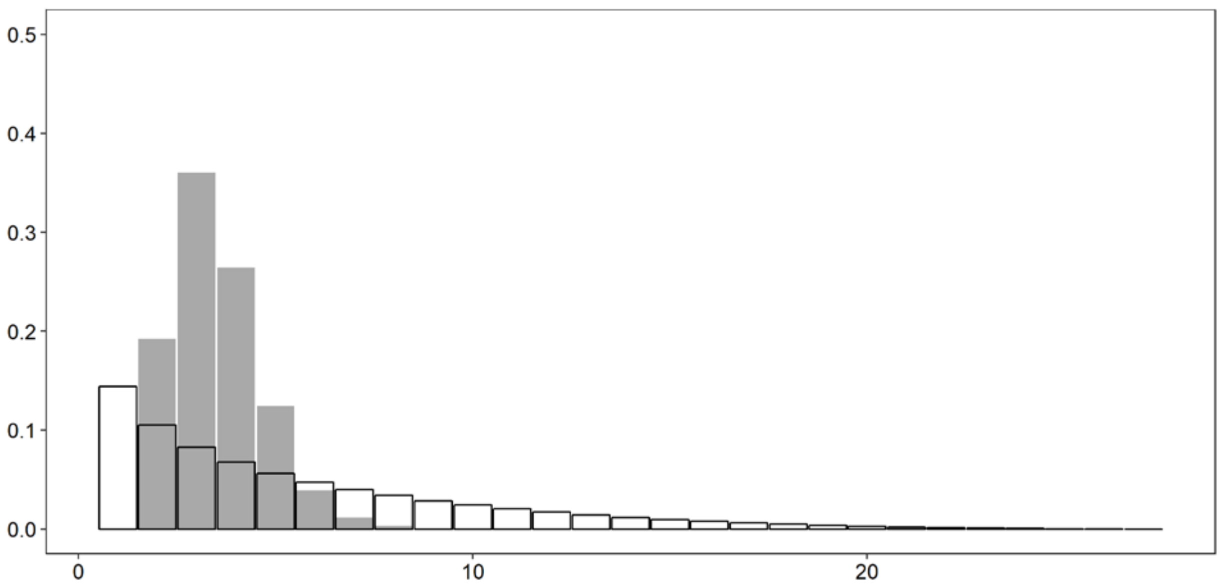


Notes: Panel 1 shows the initial 1000 replications of the algorithm and shows how it converges to a mean slightly below 4 variables. It does not converge to one single model which shows fundamental model uncertainty. Panel 2 compares the prior on the model space (empty bars) with the posterior produced by the algorithm (grey bars). The prior is based on a random inclusion probability with mean model size four as introduced in section 4.2. The posterior peaks at a model size of 3-4 variables and looks reasonably gaussian. Its mode is not the smallest model, thus we are able to update the prior towards the existence of a small model as Ley and Steel (2009) intended. All results are based on a run of the MC³ algorithm with structural reforms of employment protection legislation in regular contracts on the left hand side of the logit model.

Panel 3. Convergence of the MC^3 algorithm: Product Market Reforms

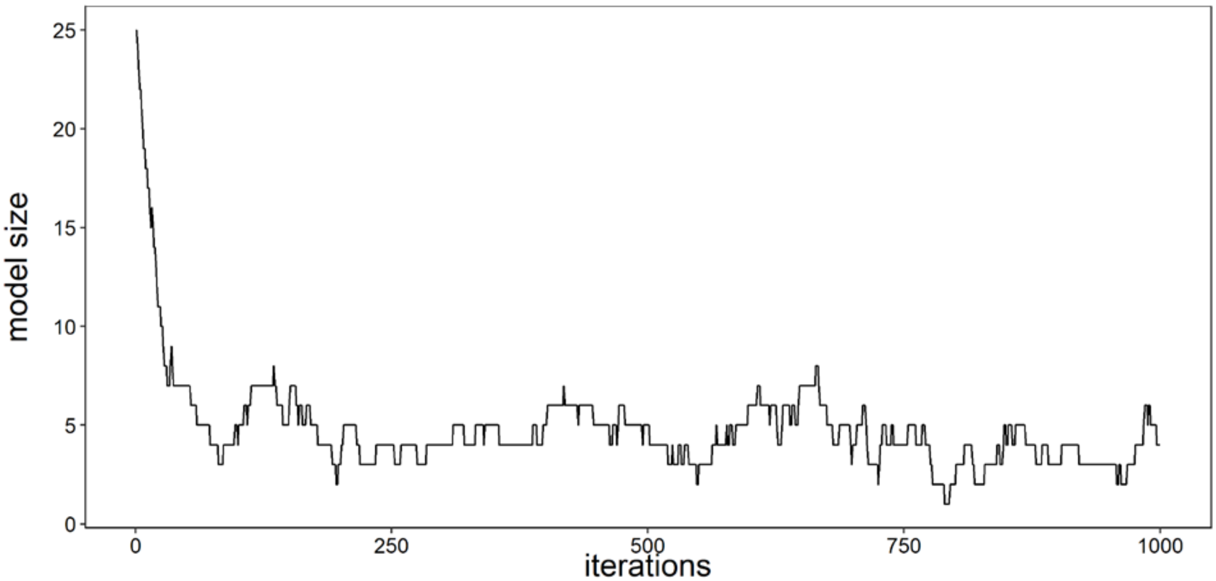


Panel 4. Prior and Posterior distribution over model size: Product Market Reforms

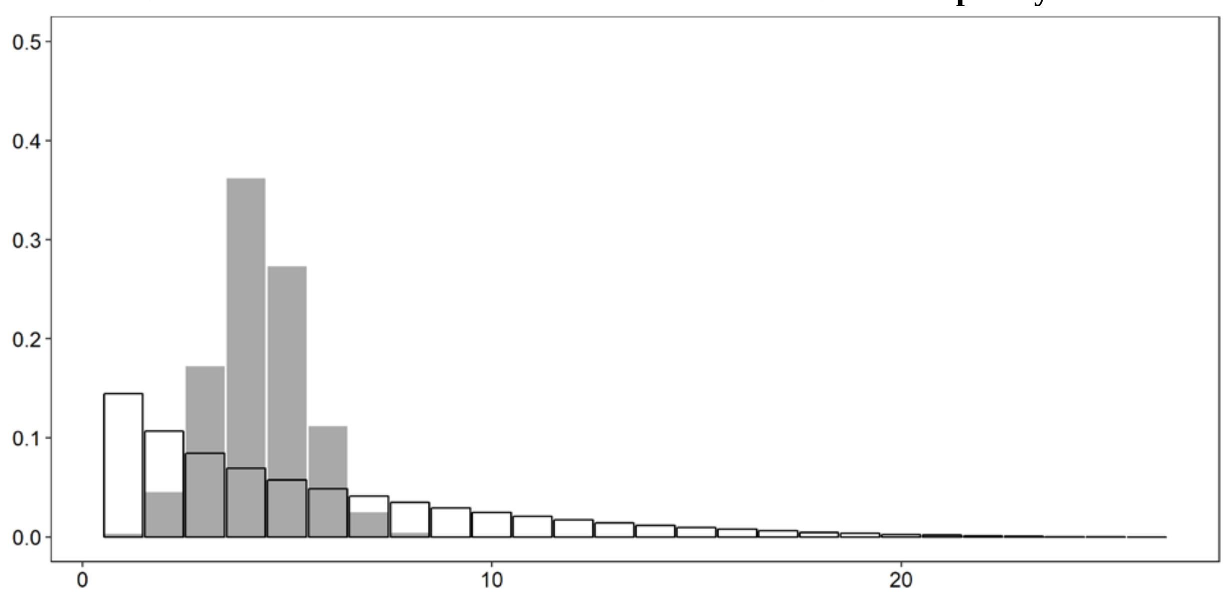


Notes: Panel 3 and 4 repeat the same exercise as panel 1 and 2 for reforms in product markets. The posterior again peaks at a model size of 3-4 variables and looks reasonably gaussian.All results are based on a run of the MC^3 algorithm with structural reforms of product markets on the left hand side of the logit model.

Panel 5. Convergence of the MC^3 algorithm: EPL temporary contracts

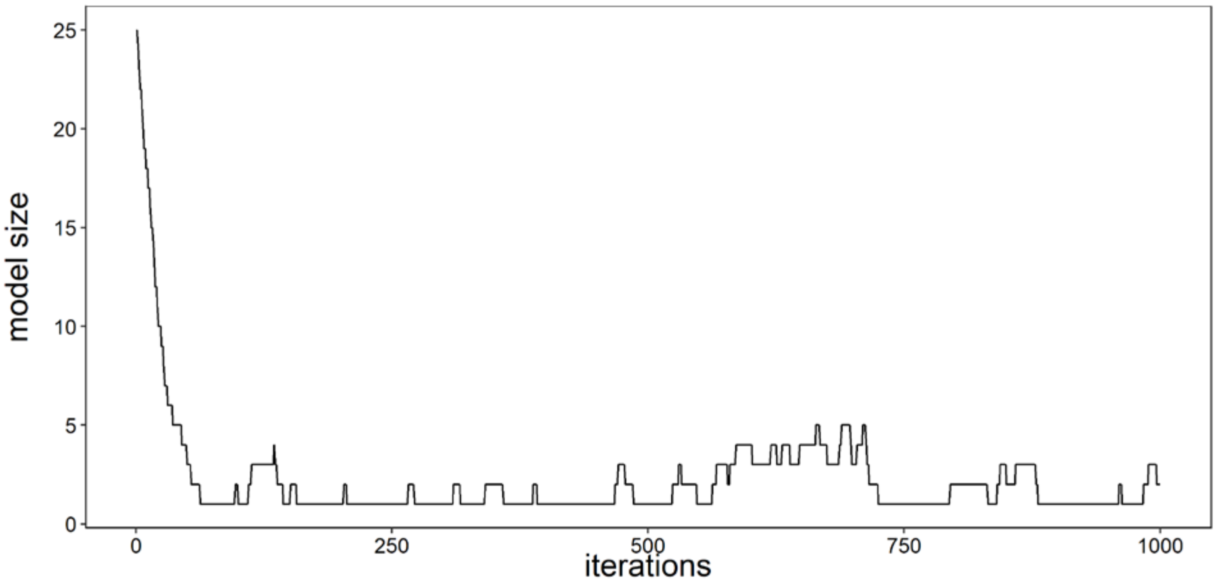


Panel 6. Prior and Posterior distribution over model size: EPL temporary contracts

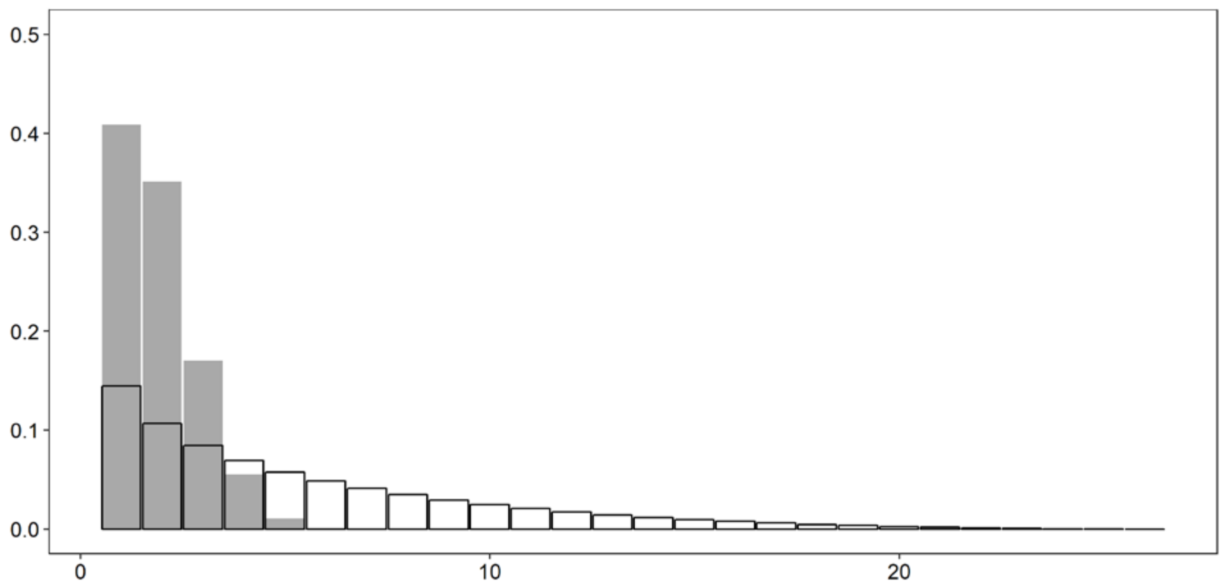


Notes: Panel 5 and 6 repeat the same exercise as panel 1 and 2 for reforms of employment protection legislation in temporary contracts. The posterior peaks at a model size of 4-5 variables and looks reasonably gaussian.All results are based on a run of the MC^3 algorithm with structural reforms of employment protection legsilation in temporary contracts on the left hand side of the logit model.

Panel 7. Convergence of the MC^3 algorithm: Unemployment Benefits



Panel 8. Prior and Posterior distribution over model size: Unemployment benefits



Notes: Panel 7 and 8 repeat the same exercise as panel 1 and 2 for reforms of unemployment benefits. Here, the posterior peaks at a model size of 1-2 variables indicating that the data is not able to distinguish well fitting models. All results are based on a run of the MC^3 algorithm with structural reforms of unemployment benefits on the left hand side of the logit model

Annex 2. BAMLE Results for Different Estimations

Table A2.1. Product market reforms

		Logit pooled			Linear Probability Model pooled			Logit time and year demeaned			Linear Probability Model time and year demeaned		
		post.mean	post.se	pip	post.mean	post.se	pip	post.mean	post.se	pip	post.mean	post.se	pip
business conditions	gdp growth	-0.488	0.341	0.069	-0.471	0.320	0.069	-0.045	0.343	0.016	-0.051	0.400	0.015
	unemployment	0.366	0.172	0.184	0.485	0.223	0.330	0.519	0.357	0.069	0.511	0.377	0.074
	deep recession	-1.575	1.707	0.037	-0.842	1.210	0.013	-0.685	0.992	0.012	-0.494	1.221	0.010
	crisis	2.499	3.304	0.038	2.764	2.524	0.031	3.692	2.861	0.040	4.213	2.963	0.056
macroeconomic policies	exchange rate	0.271	0.208	0.059	0.180	0.179	0.028	-0.350	0.365	0.025	-0.435	0.377	0.043
	taylor rate	0.106	0.218	0.027	0.045	0.255	0.016	0.103	0.300	0.017	-0.003	0.383	0.010
structural features of the domestic economy	openness	-2.439	1.533	0.093	-1.506	1.359	0.032	9.862	3.500	0.121	10.516	5.856	0.093
	old age dep.	0.477	0.317	0.100	0.445	0.293	0.063	1.138	0.521	0.089	1.171	0.646	0.051
	gini (net)	-0.007	0.185	0.024	0.044	0.222	0.012	-0.503	0.523	0.035	-0.586	0.464	0.061
	gini (market)	0.155	0.154	0.045	0.179	0.186	0.031	-0.137	0.331	0.017	-0.158	0.347	0.012
	government debt	4.609	1.919	0.293	5.394	2.351	0.341	6.332	3.707	0.061	5.031	4.228	0.054
	regulation indicator	2.181	0.653	1	1.880	0.596	1	6.671	2.068	1	7.453	1.820	1
external factors	int. spillovers	0.988	0.277	0.998	1.079	0.249	0.999						
	EU accession	9.651	6.749	0.134	10.438	5.239	0.169	5.381	4.589	0.043	8.400	5.517	0.069
	EU directives (pmr)	1.741	1.047	0.196	1.660	0.898	0.134	9.364	4.572	0.220	14.707	4.903	0.796
political factors, including ideology, structural features of the political system, and conjunctural features (e.g. political cycles and crises)	exec r.l.c (cont)	-1.157	1.034	0.050	-0.947	0.844	0.036	-1.142	1.178	0.043	-1.153	1.021	0.056
	exec left	-0.396	2.033	0.028	-1.209	1.588	0.019	-2.323	2.251	0.047	-2.261	1.959	0.055
	union density	-0.015	0.035	0.024	-0.021	0.037	0.020	0.088	0.123	0.022	0.103	0.141	0.030
	centralization gov.	1.750	2.915	0.031	1.431	2.994	0.011	1.075	5.651	0.011	1.185	5.235	0.015
	vote share gov.	-0.058	0.080	0.034	-0.052	0.060	0.028	-0.010	0.067	0.015	-0.018	0.074	0.020
	gov. controls all houses	-1.001	1.849	0.027	-1.165	1.960	0.010	-4.320	2.976	0.061	-3.138	2.469	0.050
	months to election	-0.009	0.062	0.023	-0.015	0.056	0.016	0.007	0.065	0.015	0.014	0.064	0.017
	elections next year	1.099	1.839	0.026	0.956	1.752	0.015	0.542	1.907	0.011	0.566	1.696	0.018
	years left in office	-0.512	0.667	0.029	-0.477	0.639	0.020	-0.282	0.665	0.014	-0.280	0.641	0.023
	years spent in office	0.252	0.242	0.044	0.263	0.265	0.025	0.241	0.271	0.028	0.262	0.294	0.034
momentum	domestic packaging	0.914	0.826	0.057	1.156	0.994	0.031	0.891	1.330	0.023	0.755	1.111	0.022

Note: Shows posterior means conditional on inclusion, standard errors conditional on inclusion, as well as posterior inclusion probabilities of the BAMLE exercise discussed in the methodological section over different estimators. The first three result columns repeat the baseline results based on Logit models. Columns 4-6 show the same exercise based on a linear probability model. Columns 7-9 show results of Logit models run on time and individually demeaned data.

Table A2.2. Temporary EPL reforms

		Logit pooled			Linear Probability Model pooled			Logit time and year demeaned			Linear Probability Model time and year demeaned		
		post.mean	post.se	pip	post.mean	post.se	pip	post.mean	post.se	pip	post.mean	post.se	pip
business conditions	gdp growth	-0.298	0.382	0.045	0.150	0.419	0.011	-0.844	0.427	0.100	-0.510	0.522	0.029
	unemployment	0.298	0.216	0.071	0.597	0.246	0.454	1.169	0.400	0.805	1.263	0.397	0.910
	deep recession	2.516	0.922	0.924	4.342	1.127	0.997	3.342	1.023	0.946	6.275	1.261	1.000
	crisis	1.305	2.735	0.032	0.968	2.494	0.008	1.358	2.573	0.023	2.750	2.930	0.038
macroeconomic policies	exchange rate	0.101	0.231	0.035	0.211	0.198	0.031	0.113	0.532	0.016	0.240	0.400	0.018
	taylor rate	-0.526	0.327	0.135	-0.401	0.299	0.046	-0.095	0.418	0.023	-0.246	0.473	0.021
structural features of the domestic economy	openness	-2.402	2.114	0.063	-2.854	1.731	0.071	-11.032	6.035	0.071	-3.407	6.419	0.016
	old age dep.	0.615	0.400	0.130	0.403	0.311	0.024	-0.381	0.695	0.014	-0.444	0.664	0.013
	gini (net)	0.015	0.216	0.029	0.253	0.213	0.034	0.538	0.568	0.028	0.843	0.516	0.136
	gini (market)	0.253	0.247	0.066	0.258	0.190	0.059	0.216	0.410	0.021	0.682	0.372	0.139
	government debt	2.205	2.330	0.055	2.739	2.494	0.030	-4.534	4.378	0.047	-5.727	4.694	0.044
	regulation indicator	2.670	0.645	1	2.899	0.624	1	3.945	1.831	1	3.708	1.457	1
external factors	int. spillovers	0.902	0.389	0.696	1.027	0.384	0.689						
	EU accession	3.305	6.260	0.033	2.591	5.077	0.015	2.464	4.540	0.016	4.968	5.337	0.036
political factors, including ideology, structural features of the political system, and conjunctural features (e.g. political cycles and crises)	exec r.l.c (cont)	-0.158	0.947	0.023	-0.229	0.925	0.009	-0.500	1.053	0.023	-0.130	1.111	0.028
	exec left	-0.607	1.763	0.028	-0.511	1.714	0.013	-1.361	2.038	0.035	-2.221	2.059	0.042
	union density	0.001	0.041	0.028	-0.005	0.041	0.009	0.062	0.179	0.017	-0.113	0.174	0.041
	centralization gov.	2.534	3.206	0.048	3.835	3.161	0.037	5.405	7.408	0.024	5.755	5.527	0.026
	vote share gov.	-0.074	0.046	0.063	-0.064	0.061	0.019	-0.055	0.048	0.027	-0.061	0.074	0.037
	gov. controls all houses	-0.551	1.936	0.028	0.497	2.060	0.009	1.086	2.608	0.021	-0.098	2.474	0.013
	months to election	0.002	0.062	0.027	0.001	0.059	0.011	-0.005	0.069	0.009	-0.009	0.059	0.022
	elections next year	-0.892	1.833	0.030	-0.440	1.810	0.013	-0.417	1.830	0.012	-0.322	1.705	0.018
	years left in office	0.470	0.754	0.032	0.436	0.684	0.012	0.285	0.816	0.017	0.268	0.671	0.025
	years spent in office	-0.057	0.260	0.022	0.053	0.280	0.014	0.235	0.297	0.024	0.230	0.308	0.034
momentum	domestic packaging	3.235	1.054	0.968	4.067	1.129	0.965	2.883	1.534	0.305	3.104	1.300	0.457

Note: Shows posterior means conditional on inclusion, standard errors conditional on inclusion, as well as posterior inclusion probabilities of the BAMLE exercise discussed in the methodological section over different estimators. The first three result columns repeat the baseline results based on Logit models. Columns 4-6 show the same exercise based on a linear probability model. Columns 7-9 show results of Logit models run on time and individually demeaned data.

Table A2.3. Unemployment benefit reforms

		Logit pooled			Linear Probability Model pooled			Logit time and year demeaned			Linear Probability Model time and year demeaned		
		post.mean	post.se	pip	post.mean	post.se	pip	post.mean	post.se	pip	post.mean	post.se	pip
business conditions	gdp growth	0.323	0.262	0.059	0.291	0.264	0.038	0.160	0.248	0.021	0.032	0.350	0.019
	unemployment	0.292	0.172	0.125	0.309	0.186	0.097	0.628	0.273	0.255	0.675	0.311	0.298
	deep recession	-0.166	0.838	0.031	-0.185	0.902	0.017	0.079	0.763	0.016	0.286	1.000	0.020
	crisis	-1.184	1.699	0.035	-1.430	2.054	0.019	-0.483	2.065	0.019	-0.861	2.482	0.020
macroeconomic policies	exchange rate	0.236	0.172	0.070	0.238	0.165	0.056	-0.131	0.318	0.022	-0.204	0.316	0.028
	taylor rate	0.199	0.192	0.049	0.181	0.204	0.034	0.112	0.330	0.019	-0.020	0.316	0.016
structural features of the domestic economy	openness	1.181	0.760	0.045	1.097	1.252	0.025	1.746	5.210	0.020	4.530	4.999	0.030
	old age dep.	0.014	0.264	0.030	0.013	0.241	0.015	-0.105	0.396	0.013	0.061	0.509	0.013
	gini (net)	0.026	0.169	0.030	0.002	0.173	0.015	-0.059	0.499	0.022	-0.081	0.404	0.021
	gini (market)	0.154	0.152	0.045	0.087	0.145	0.024	0.315	0.294	0.039	0.371	0.280	0.046
	government debt	1.382	1.610	0.038	1.287	1.941	0.028	3.257	2.674	0.041	3.529	3.423	0.043
	regulation indicator	0.124	0.058	1	0.125	0.050	1	0.013	0.099	1	0.014	0.104	1
external factors	int. spillovers	0.050	0.363	0.031	0.030	0.391	0.017						
	EU accession	3.709	6.622	0.037	3.144	4.383	0.030	2.076	4.087	0.019	1.969	4.590	0.012
political factors, including ideology, structural features of the political system, and conjunctural features (e.g. political cycles and crises)	exec r.l.c (cont)	0.711	0.839	0.034	0.698	0.880	0.022	0.910	1.059	0.031	0.867	1.041	0.032
	exec left	-1.074	1.817	0.032	-0.843	1.572	0.024	-0.742	2.098	0.024	-1.027	2.396	0.025
	union density	0.024	0.029	0.041	0.027	0.032	0.028	0.163	0.070	0.067	0.133	0.111	0.047
	centralization gov.	-2.270	2.808	0.040	-2.299	2.542	0.029	-3.941	4.376	0.031	-3.892	4.300	0.033
	vote share gov.	-0.006	0.059	0.029	-0.003	0.048	0.022	-0.001	0.054	0.022	-0.007	0.059	0.019
	gov. controls all houses	0.032	1.852	0.029	-0.162	1.623	0.017	1.218	1.622	0.025	1.164	2.082	0.024
	months to election	0.010	0.049	0.029	0.009	0.057	0.016	-0.009	0.056	0.014	-0.002	0.054	0.018
	elections next year	0.346	1.477	0.030	0.361	1.454	0.020	0.445	1.531	0.019	0.358	1.386	0.018
	years left in office	0.039	0.496	0.029	0.054	0.571	0.018	0.025	0.596	0.022	0.061	0.541	0.020
	years spent in office	-0.207	0.244	0.044	-0.193	0.217	0.032	-0.529	0.361	0.147	-0.425	0.243	0.111
momentum	domestic packaging	0.168	0.679	0.030	0.117	0.704	0.020	0.452	1.050	0.027	0.427	0.881	0.026

Note: Shows posterior means conditional on inclusion, standard errors conditional on inclusion, as well as posterior inclusion probabilities of the BAMLE exercise discussed in the methodological section over different estimators. The first three result columns repeat the baseline results based on Logit models. Columns 4-6 show the same exercise based on a linear probability model. Columns 7-9 show results of Logit models run on time and individually demeaned data.

Annex 3. Robustness of Results to the Choice of Priors

Table A3.1. Robustness of results to extreme prior choices

Product Market Reforms

Panel 1, prior: $m = 1$	pip	Panel 2, prior: $m = K/2$	pip	Panel 3, prior: $m = K$	pip
regulation indicator	1	regulation indicator	1	regulation indicator	1
int. spillovers	1.000	int. spillovers	1.000	int. spillovers	1.000
government debt	0.224	government debt	0.372	EU directives (pmr)	0.596
unemployment	0.148	EU directives (pmr)	0.289	government debt	0.509
EU directives (pmr)	0.137	unemployment	0.236	unemployment	0.385
EU accession	0.092	EU accession	0.148	exchange rate	0.317
old age dep.	0.084	openness	0.118	EU accession	0.301
openness	0.063	exchange rate	0.113	gdp growth	0.209
gdp growth	0.058	gdp growth	0.088	old age dep.	0.161
exchange rate	0.051	old age dep.	0.072	openness	0.144

EPL temporary contracts

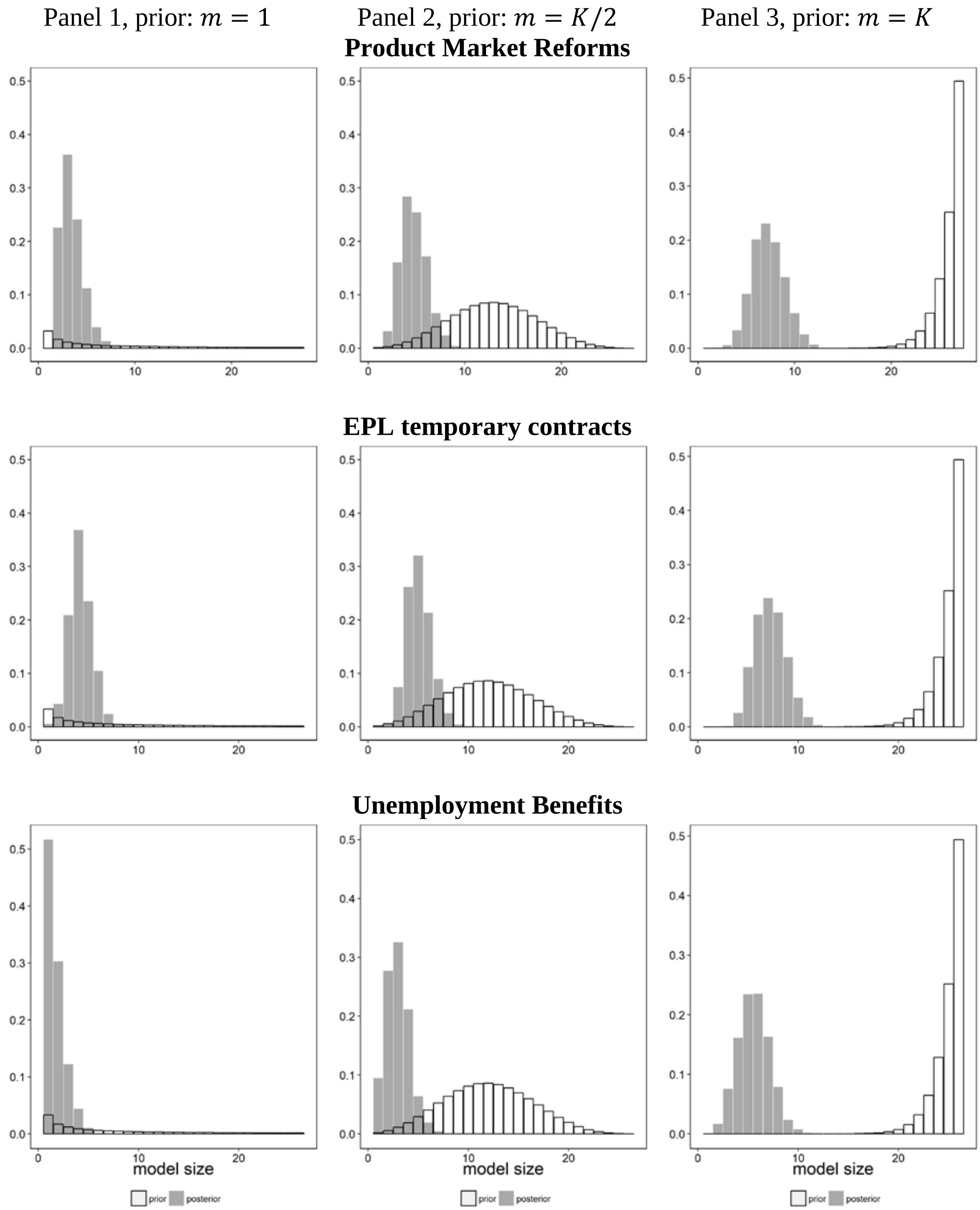
Panel 1, prior: $m = 1$	pip	Panel 2, prior: $m = K/2$	pip	Panel 3, prior: $m = K$	pip
regulation indicator	1	regulation indicator	1	regulation indicator	1
domestic packaging	0.970	domestic packaging	0.986	domestic packaging	0.999
deep recession	0.919	deep recession	0.966	deep recession	0.996
int. spillovers	0.651	int. spillovers	0.789	int. spillovers	0.930
taylor rate	0.108	taylor rate	0.127	taylor rate	0.207
old age dep.	0.086	old age dep.	0.086	old age dep.	0.173
unemployment	0.055	vote share gov.	0.077	vote share gov.	0.122
vote share gov.	0.039	unemployment	0.072	gini (market)	0.119
gini (market)	0.037	gini (market)	0.067	openness	0.116
openness	0.037	openness	0.051	exchange rate	0.092

Unemployment Benefits

Panel 1, prior: $m = 1$	pip	Panel 2, prior: $m = K/2$	pip	Panel 3, prior: $m = K$	pip
regulation indicator	1	regulation indicator	1	regulation indicator	1
gdp growth	0.032	gdp growth	0.107	gdp growth	0.171
unemployment	0.075	unemployment	0.223	unemployment	0.569
deep recession	0.013	deep recession	0.039	deep recession	0.062
crisis	0.015	crisis	0.035	crisis	0.084
exchange rate	0.034	exchange rate	0.115	exchange rate	0.396
taylor rate	0.027	taylor rate	0.080	taylor rate	0.099
openness	0.020	openness	0.067	openness	0.225
old age dep.	0.013	old age dep.	0.036	old age dep.	0.075
gini (net)	0.012	gini (net)	0.036	gini (net)	0.044

Notes: Each panel shows the 10 variables with the highest posterior inclusion probability (pip) as well as their pip for different choices of the hyperparameter m. In the first panel, we choose the lower extreme (prior: $m = 1$). In the second panel, we assume that a model with a model size of half of the available variables is the most likely model (prior: $m = K/2$). In the third panel, we consider the maximum number of variables as the most likely model (prior: m=K).

Figure A3.1. Robustness of posterior model size to extreme prior choices



Notes: These figures show results in line with Figure 7 for the other three reform areas. The barcharts show prior (empty framed bars) and posterior (grey bars) model sizes for three different choices of hyperparameter m. In panel 1, we choose the lower extreme (prior: $m = 1$), in panel 2, half the available variables (Prior: $m = K/2$) and in panel 3 the maximum number of variables. Results are based on reforms in product markets (first row), EPL in temporary contracts (second row) and reforms of unemployment benefits (third row).