Heterogeneous Gravity: Democracy and Bilateral Trade Flows*

Rodolphe Desbordes[1,2], Markus Eberhardt[3,4], and Mario Larch[5]

[1]SKEMA Business School Paris, France
[2]Université Côte d'Azur, France
[3]School of Economics, University of Nottingham, U.K.
[4]Centre for Economic Policy Research, London, U.K.
[5]University of Bayreuth, Germany

November 8, 2024

Abstract: We study the effect of democratic regime change on bilateral trade, extending standard gravity empirics to 'heterogeneous gravity' estimated at the country-pair level. Our difference-in-differences implementation accounts for multilateral resistance and selection into regime change. We find treatment effects of 40% higher exports for countries with thirty years in democracy, though effects vary by destination regime status. Split sample analysis reveals insignificant effects before 1989, after the Cold War average effects remain below 20%. Comparing favourable and unfavourable geography, legal origin, culture, and history we establish that exporters with unfavourable 'deep determinants' commonly experience lower export effects from regime change.

Keywords: trade gravity model, heterogeneity, democratic regime change, panel data, interactive fixed effects, deep determinants of comparative development

JEL codes: F13, F14, P16, C23

*We thank Arturas Juodis for comments on the empirical implementation in an earlier draft and participants at the Nottingham GEP brown bag seminar for useful comments and suggestions. The usual disclaimers apply.
Correspondence: Markus Eberhardt, School of Economics, Sir Clive Granger Building, University of Nottingham, University Park, Nottingham NG7 2RD, U.K. Email: markus.eberhardt@nottingham.ac.uk.

1 Introduction

Economists generally agree on the importance of good institutions for economic prosperity, and the causal effect of becoming a democracy — a 'bundle' of institutions — on long-run income per capita is suggested to be on the order of 20-30% (Acemoglu et al. 2019, Boese-Schlosser & Eberhardt 2024).[1] Similarly, the relevance of trade for economic development is widely recognised among economists and policymakers alike: trade flows are associated with efficiency gains, technology transfer, and increased innovation activities (Costinot & Rodríguez-Clare 2014, Feyrer 2019). To the uninitiated, it may therefore come as a surprise that the literature on democratic institutions in their effect on bilateral trade flows — using the workhorse 'gravity model' approach[2] — is quite sparse. They might also notice that other important *country-specific* characteristics or policy levers countries may employ to become more integrated with the global economy, such as tariffs, exchange rate arrangements, or corporate taxation, are also largely ignored by trade gravity analysis.

Counter-intuitively, the solution to this puzzle relates to the significant progress made in the literature: the structural gravity model has been considerably strengthened by theoretical work deriving it from a wide class of microeconomic principles (Arkolakis et al. 2012, Costinot & Rodríguez-Clare 2014, Allen et al. 2020), and further by contributions to the empirical methodology which prescribe what the gravity specification should look like, and how it should be estimated to address a range of econometric challenges, including the global network aspect of trade ('multilateral resistance') as well as concerns over heteroskedastic residuals and zero trade flows (Anderson & Van Wincoop 2003, Santos Silva & Tenreyro 2006, Baier & Bergstrand 2007). The empirical path taken by the gravity literature to achieve credibility and rigour is also its demonstrable weakness, an empirical straitjacket of sorts: identification is achieved by focusing on the narrow set of *dyadic* policy instruments, with the analysis of country-specific ones largely ignored by the 'diktat' of the structural gravity model specification. Exceptions include Heid et al. (2021) and Beverelli et al. (2024), who seek alternative means to identify country-specific effects within the structural gravity model, involving intra-national sales, albeit for limited time series. The merit of this analysis hinges on the assumption that regime change affects domestic and international sales differentially and that quantifying this *difference* is economically meaningful, rather than establishing whether 'democracy does cause higher trade' (like we do). Other work has sought to construct dyadic democracy variables to get around the identification problem (e.g. Yu 2010, Álvarez et al. 2018) but cannot speak to the impact of

[1]These studies counter widespread scepticism among economists and political scientists about the democracy-growth nexus. There is the suggestion that allowing voters to remove an incumbent government through the power of the electoral process would drive up consumption and reduce the investment rate to the detriment of economic growth (e.g. Baum & Lake 2003, 334f). Studying China or the East Asian 'Tiger' economies suggests that democracy is not a necessary condition for economic prosperity, but this form of cherry-picking of autocratic success stories ignores the established link between autocracy and the large variation in economic performance (Persson & Tabellini 2009, Knutsen 2012, Imam & Temple 2024) .

[2]For recent surveys see Head & Mayer (2014), Yotov et al. (2016), and Yotov (2024).

national (rather than bilateral) democratic regime change. A substantial number of papers simply ignore the 'prescribed' set of fixed effects in their analysis of monadic variable (e.g. Dutt & Traca 2010, Francois & Manchin 2013). Finally, some studies (e.g. Eaton & Kortum 2002, Head & Ries 2008) argue that country-specific effects could be extracted in a second-step regression from the fixed effects of a structural gravity model. This approach has been undermined by the finding that in the Poisson Pseudo Maximum Likelihood (PPML) model favoured by the empirical literature said fixed effects are perfect matches for multilateral resistance terms and hence any impact of additional covariates in the second step is unidentified (Fally 2015, Beverelli et al. 2024).

Theoretical considerations regarding a democracy-trade nexus build on institutional notions of property rights protection and fair competition on the exporter side, implying that exporting firms are enabled to develop higher-quality products, which then naturally create a global demand (Yu 2010). On the importer side, improving democratic institutions is associated with a shift to pro-trade policies in developing countries, where power shifts away from protectionist political and economic elites (Milner & Kubota 2005) and the reverse in advanced economies (O'Rourke & Taylor 2006), echoing the widespread suggestion that the median voter is poor and favours consumption over investment (e.g. Przeworski et al. 2000), hence suggesting an ambiguous effect of democracy on trade.[3]

In this paper, we propose an empirical approach to break the empirical straitjacket: we investigate the gravity relationship at the country-pair level, account for the manifestation of the trade network as well as regime change endogeneity, and estimate the model in its multiplicative form, enabling us to investigate the effect of *monadic* variables such as democracy on trade flows. We adopt a multifactor error structure to capture multilateral resistance and other forms of unobserved time-varying heterogeneity, treating the factors as nuisance terms in the same fashion as the standard pooled gravity model employs a myriad of fixed effects. Following Boneva & Linton (2017) we use cross-section averages to proxy common factors (an approach pioneered by Pesaran 2006), in combination with the spirit of a heterogeneous treatment effects estimator by Chan & Kwok (2022):[4] we construct the cross-section averages from the sample of 'never-treated' countries ('never-democracies') and test whether the 'information' captured is equally relevant for the treated sample.[5] We estimate a treatment equation for the bilateral trade flows of countries that experienced regime change during the sample period. The empirical implementation is via PPML at the pair-level to

[3]The latter argument of a diversion from investment to consumption is similarly applied in the democracy-growth context (see footnote 1) yet the empirical literature has found resoundingly positive and large causal effects (Acemoglu et al. 2019, Eberhardt 2022, Boese-Schlosser & Eberhardt 2024).

[4]In the democracy-growth literature economists favour the use of binary indicators for democracy versus autocracy (Papaioannou & Siourounis 2008, Acemoglu et al. 2019, Boese-Schlosser & Eberhardt 2023, 2024), which we construct from V-Dem data. Democratic regime change is endogenous given the link between economic integration and institutional change found in the literature (e.g. Acemoglu et al. 2005, Nunn & Trefler 2014, Puga & Trefler 2014).

[5]Pooled difference-in-differences estimators rely on the assumption of parallel trends before treatment. The Chan & Kwok (2022) estimator can allow for non-parallel trends but appeals to the weaker assumption of equal average factor loadings on the common factors between treated and control samples — see Section 2 for details.

address concerns over heteroskedastic errors, in line with the existing best practice in the literature. We need a further correction to address the practical issue of country-pairs where the exporter or importer experienced regime change but no trade *ever* takes place between the countries.[6]

We study the export flows of 84 'treated' countries which during 1950-2014 experienced 111 democratic regime changes and 55 reversals to autocracy. Presenting our results in 5-year treatment intervals, we find that the economic effects of democratic regime change are sizeable: the average treatment effect after twenty-five (thirty) years in democracy amounts to a 30% (40%) rise in exports. However, in line with recent results in the democracy-growth literature, export effects are lacklustre in the first decade and a half after regime change: insignificant and at times even negative-significant.[7] Our results show differences across *importer* regime types, whereby the variance of the export effect for *autocratic* destinations is substantially higher and hence frequently statistically insignificant. We conduct a range of robustness checks including alternative definitions of regime change and focusing on countries with just a single regime change, but the patterns remain qualitatively unchanged. Splitting the dataset into the Cold War and post-Cold War eras shows that democratic regime change had no significant trade effect in the former, whereas the analysis of 1990-2014 follows similar patterns to the full sample results but effects peak at below 20%. Finally, our analysis of deep determinants of comparative development establishes the unequal economic benefits of regime change: countries with unfavourable deep determinants (e.g. few frost days, no European descendants) commonly experienced lower export effects per year in democracy than their peers, with the significance of culture emerging strongly in the Post-Cold War period.

We make three contributions to the literature: first, we introduce a simple but powerful new empirical implementation of the gravity model of trade and other flows which allows for the identification of monadic variables. Second, we study the relationship between country-specific democratic regime change and bilateral trade flows, contributing to the investigation of underlying transmission mechanisms of the sizeable democracy-growth effect. Third, we link our heterogeneity analysis of unequal benefits from democratic regime change to the work on 'deep determinants of comparative development' to demonstrate that such immutable structural characteristics prevail and can substantially diminish the large positive economic effect of democracy.

2 Methodology

In this Section we introduce our empirical estimator. We begin by detailing its five important elements and the literature(s) these are taken from. We then explain our approach in plain words

[6] We cannot estimate a pairwise regression in this case yet simply ignoring such pairs would induce selection bias (of pairs with positive trade effects at the intensive or extensive margin).

[7] These patterns can be rationalised by the upheaval following regime change and/or 'democratic overload' (Gerring et al. 2005, Cervellati & Sunde 2014, Boese-Schlosser & Eberhardt 2024).

before introducing the estimation equation. A more detailed exposition is provided in Appendix E. The final section details an identification test adjusted from Chan & Kwok (2022).

2.1 A Multi-Faceted Implementation

Our difference-in-differences implementation to capture binary monadic variables in a factor-augmented heterogeneous gravity model has the following conceptual ingredients: (i) We build on the literature for heterogeneous parameter models (Swamy 1970, Pesaran & Smith 1995) and estimate the gravity model at the pair-level. (ii) We can do so because we treat multilateral resistance as unobserved time-varying heterogeneity and adopt a common factor framework to capture this in the model (Pesaran 2006, Bai 2009, Baier & Standaert 2024). (iii) We follow the gravity literature in estimating a generalised linear (PPML) model to address concerns over heteroskedastic errors and zero trade flows (Santos Silva & Tenreyro 2006), but introduce a variant estimated at the pair-level and embedded in a common factor framework using cross-section averages (CA) as factor proxies (Boneva & Linton 2017). (iv) Specifying democracy as a binary 'treatment' variable, we borrow from the treatment effects literature adopting common factors (Gobillon & Magnac 2016, Xu 2017, Chan & Kwok 2022) and estimate the gravity model at the pair-level for exporters i or destinations j which experienced regime change during the sample period (treated sample). We further use the 'never-treated' exporters and importers (the control sample) to construct the factor proxies (cross-section averages), for causal identification of endogenous regime change, mimicking the Chan & Kwok (2022) strategy. We further adapt their 'weak parallel trend test' to the three-way panel context to test the identifying assumption of expected factor-loading equality between treated and control samples. (v) Finally, our pairwise regression is not identified for treated exporters or importers if trade flows to some destinations remain zero in all time periods. We address this selection problem by including factor proxies from the zero-trade pairs in the treatment regression.

We refer to this as the heterogeneous PPML-CCE-DID estimator to highlight its most significant 'origins' (Santos Silva & Tenreyro 2006, Pesaran 2006, Chan & Kwok 2022, respectively).

2.2 A Narrative Explanation of the Proposed Estimator

Our estimator compares trade flows between exporters (destinations) that experienced regime change with those that did not, hence it is a Difference-in-Differences implementation. We adopt the ideas of Chan & Kwok (2022) who propose a heterogeneous Difference-in-Differences estimator and thus obtain treatment effect estimates at the pair level. In line with their empirical strategy, the regressions are only run for treated exporters (destinations) and the treatment equation is augmented with proxies for unobserved common factors constructed from the never-treated samples (exporters *and* destinations). As these factors have heterogeneous parameters they can capture a great deal of unobserved time-varying heterogeneity, which explains the popularity of the common factor framework in the cross-country empirical literature (e.g. Eberhardt et al. 2013, Eberhardt & Presbitero 2015, Boese-Schlosser & Eberhardt 2023, 2024). Chan & Kwok (2022) follow Bai (2009) in adopting a principal component approach to extracting factors from the residuals of a control sample regression. In our three-way panel we instead employ the Pesaran (2006) approach of using cross-section averages (CA) as factor proxies instead.[8] We chose the standard 'economic mass' variables (GDP, population) for exporters and importers to construct these CA. From Boneva & Linton (2017) we know that we can apply the same CA-augmentation in a generalised linear model such as a logit or Poisson model and so we implement the CCE-DID model described in a PPML regression.[9]

2.3 Estimation and Inference

Using PPML, we estimate the following equation at the pair-level for exporters i and destinations j which experienced democratic regime change:

$$\begin{aligned} \mu_{ijt} \quad = \quad & exp[\alpha_{ij} + \gamma^1_{ij}\,\text{FTA}_{ijt} + \gamma^2_{ij}\,\text{Common Currency}_{ijt} + \theta_{ij} D_{it} + \eta_{ij} D_{jt} \\ & +\delta^i_{ij}\,\overline{ln(Y)}^i_t + \kappa^i_{ij}\,\overline{ln(Pop)}^i_t + \delta^j_{ij}\,\overline{ln(Y)}^j_t + \kappa^j_{ij}\,\overline{ln(Pop)}^j_t \\ & +\delta^{i0}_{ij}\,\overline{ln(Y)}^{i0}_t + \kappa^{i0}_{ij}\,\overline{ln(Pop)}^{i0}_t + \delta^{j0}_{ij}\,\overline{ln(Y)}^{j0}_t + \kappa^{j0}_{ij}\,\overline{ln(Pop)}^{j0}_t + \epsilon_{ijt}] \qquad \forall ij,\, j \neq i. \end{aligned} \tag{1}$$

The dependent variable are the exports from i to j at time t. Our parameters of interest are θ_{ij} and η_{ij} which relate to the regime change dummies for exporters D_{it} and destinations D_{jt}. FTA and Common Currency are dyadic controls. The second line of equation (1) represents the cross-section averages at time t ($\overline{ln(X)}^i_t = N^{-1}\sum_i ln(X)_{it}$) of the 'economic mass' variables employed in the gravity model (GDP, population; both in logarithms).[10] These terms are computed from respectively all exporters i (first two terms) and all destinations j (second two terms) which remained autocratic. The final row includes the error term and four more cross-section averages, computed from GDP and population data for exporters i and destinations j which experienced regime change but for which trade flows to some destinations remained zero.

This specification is very demanding, with 13 parameters to be estimated in each country-pair regression. We average treatment estimates $\hat{\theta}_{ij}$ to yield ATETs (see Appendix Table D-1) or relate them to treatment length (see Figure 1 and Table 1), using an M-estimator (Rousseeuw & Leroy 2005) to reduce the effect of outliers — these are Mean Group estimates: $\hat{\theta}^{MG} = N^{-1}\sum_i \omega_{ij}\hat{\theta}_{ij}$ (Pesaran &

[8]Using the Bai (2009) approach would be computationally and practically infeasible since factors are extracted from exporter and destination control samples, in an unbalanced panel (requiring an expectation maximisation algorithm), with uncertainty over the number of factors (principal components) to extract.

[9]Under the assumption that the adopted factor proxies span the space of the unobserved factors.

[10]Like in the 'gold standard' pooled gravity regression where these economic mass terms for exporters and importers are not identified in the presence of exporter-time and importer-time fixed effects, they are not identified in our country-pair regression due to the presence of the CA terms.

Smith 1995) with granular weights ω_{ij} and N the number of pairwise estimates. Our nonparametric variance estimator follows Pesaran (2006): $\widehat{\text{var}}(\hat{\theta}^{MG}) = [N(N-1)]^{-1} \sum_{i=1}^{N} (\hat{\theta}_{ij} - \hat{\theta}^{MG})^2$.

2.4 Testing identifying assumptions

The reliability of every difference-in-differences estimator lives and dies by the validity of the 'parallel trend' assumption: if the soon-to-be-treated unit was already on a different trajectory to the control unit(s) before treatment, then the standard difference-in-differences estimates are biased. In the Chan & Kwok (2022) PCDID case the estimator allows for non-parallel trends across panel units, most importantly between those in the treated and those in the control samples. Nevertheless, the estimator requires an equivalent of the parallel trend test for validity, albeit in weaker form: the average factor loadings on the common factors (extracted from the control sample) in the treatment equation have to be equal those in the control equation. In practice, this implies that the estimated parameters on the factor proxies in the treatment equation statistically should not differ from 1. The intuition of this test is to ask whether the information we extract from the control sample is equally relevant in the treatment sample: imagine a world in which democratic regime change only happens in rich countries, whereas poor countries always stay autocratic. The 'information' contained in the factor proxies constructed from the poor-country control sample is then likely to be quite uninformative about the unobservables driving outcomes in the rich-country treatment sample: countries at different ends of the income scale have different economic structures, financial systems, physical and human capital stocks, etc. — we'd be proxying apples with oranges. Passing the Alpha test assures us that the unobservables are *on average the same* across treatment and control samples, that we are adopting apples to proxy for apples, or oranges for oranges.

In comparison to the Chan and Kwok (2022) Alpha test, in our setup using cross-section averages we cannot use their testing strategy, which employs the cross-section average of the residuals from the auxiliary regressions in the control sample. We devise a 'Pseudo Alpha' test for the heterogeneous trade flow regressions augmented with cross-section averages by testing whether the averaged (Mean Group) coefficients on the CA, $\bar{\hat{\delta}}^i$, $\bar{\hat{\delta}}^j$, $\bar{\hat{\kappa}}^i$ and $\bar{\hat{\kappa}}^j$ from equation (1), will jointly be equal to 1. If the underlying null of equal average factor loadings between treatment and control samples is rejected, the treatment regression may be misspecified and hence deliver biased treatment estimates.

3 Data Sources and Transformations

3.1 Gravity Model Variables

We use bilateral trade flows from TRADHIST (Fouquin & Hugot 2016, version 4) for trade flows, exporter and importer GDP and Population, combined with CEPII data for FTAs and currency unions.[11] For the export flow data we code the 'plausible zeros' as zero rather than missing. Our coverage is 1950-2014.[12] Data on democracy is taken from the Varieties of Democracy (V-Dem) project (Coppedge et al. 2021): we primarily focus on their liberal democracy index but also explore the polyarchy and liberal component indices. These data are described in great detail in Boese (2019). In a robustness check, we adopt the definition of democracy by Acemoglu et al. (2019). We discard country-pairs with fewer than 14 observations over time — this is by the necessity of the demands of our heterogeneous panel estimator.

We follow Boese-Schlosser & Eberhardt (2023) and adopt the full gravity data sample mean for liberal democracy (0.34) over 1950-2014 as the threshold for democratic regime change. The liberal democracy sample includes 822,946 observations for 17,291 country pairs (average $T = 48$). Of these 8,568 pairs are for regime changes in the exporter country, relating to 84 countries which have between 15 and 164 trading partners.[13] In robustness checks, we add or subtract either a quarter or half a sample standard deviation to this mean index value to adopt a tighter/looser definition of regime change. For our baseline specification using Liberal Democracy we provide detailed distributions of regime change events and total time spent in democracy in Appendix A. The latter suggests that the thousands of democracy estimates making up the results we present below are relatively uniformly distributed across these different periods in democracy.

3.2 Proxies for Deep Determinants

For the heterogeneity analysis we adopt a range of measures as proxies for the 'deep determinants' of comparative development, relating to climate and historical disease environment (geography), legal origin, culture (individualism, language similarity, European settlers), and colonial history. Sources include Murray & Schaller (2010), Nunn & Puga (2012), Spolaore & Wacziarg (2013), Gorodnichenko & Roland (2017) and Becker (2019). We provide more details in Appendix A.

Our empirical setup builds on crude sample splits of deep determinants into 'favourable' and 'unfavourable' for economic development informed by the existing literature. For a range of proxies, this is straightforward, e.g. French legal origin. For the proxies which are indices/continuous, we select the full exporter country sample median to split the sample.

[11] All monetary values are in nominal British Pounds.

[12] Earlier years are available, but the unbalanced country coverage creates worse coverage for countries which remained autocracies *throughout* the sample period. This in turn affects the ability of the cross-section averages to capture unobserved common factors in these early years and leads to the failure of the weak parallel trend test we devise.

[13] There are a total of 169 destinations in this baseline model. For the analysis of regime change *in a destination*, Figure 1, panel (c), there are 8,436 pairs, relating to 84 countries.

4 Results

4.1 Main Results and Robustness Checks

Main Results Figure 1 presents the treatment effects by length of time in democracy and Appendix Table D-1 the equivalent ATETs (ignoring treatment length) as well as Pseudo-Alpha test results. We use an M-estimator (Rousseeuw & Leroy 2005) to compute outlier-robust means, and all specifications (unless discussed) satisfy the Pseudo-Alpha specification test. In each plot we report mean estimates (diamonds), associated 95% confidence intervals (CI, coloured bars), and indicate the number of country-pairs for each estimate. Unless imndicated, these are average treatment estimates for exporters, using the specification with controls. We report full sample results and distinguish by trade partner/destination regime status (always autocratic or always democratic).[14]

Focusing on the full sample results (with blue CI) in Figure 1 panel (a) we can see volatile effects in the first 15 years, followed by around 20%, 30%, and 40% higher exports for countries which spent 16-20, 21-25 and 26-30 years in democracy, respectively.[15] The effect for countries which have been democratic for over 30 years is attenuated, just over 20%. Mean estimates for the subsample of results where destinations are always democratic (see-through CI) closely match the pattern described, whereas for the subsample of autocratic destinations (teal-coloured CI) we observe higher variances, which leads to mostly statistically insignificant results.

These baseline results ignore that some countries experienced *repeated* regime change and reversal to autocracy. In panel (b) we restrict results to exporters with just a single democratisation (the dominant experience) and find largely consistent early-year patterns. Countries with over 30 years in democracy now enjoy 75% higher exports, and estimates for other treatment lengths are also somewhat larger.[16]

In panel (c) we report the effects on exporter trade flows of *destination* countries experiencing regime change (CI in gold), with panel (d) again restricted to destinations with a single regime change. The subsample results are now for *exporters* which are always democratic (CI in grey) or always autocratic (CI in pink). Results are qualitatively very similar to before, albeit with the absence of a negative effect for countries with 11-15 years in democracy.

In panels (e) and (f) we return to analysing exporter regime change and present the results

[14] In Appendix Figure B-1 we include estimates for destinations which themselves experienced regime change. We exclude these in the main text for the conceptual reason that these conflate the effects when destinations are autocracies and democracies in a non-trivial manner and for the practical reason that their variability hampers ease of illustration.

[15] As is found in the existing literature on democracy and growth (Acemoglu et al. 2019, Boese-Schlosser & Eberhardt 2023), the initial period following regime change is accompanied by economic 'hardship' or at least relative stagnation, in the sense that the causal effect of democratic regime change is negative and at times statistically significant. Other than the 'tumultuous youth' experienced by many democracies (Gerring et al. 2005), at least some of these detrimental effects may be due to the crude way we define democratic regime change. We highlight this in our robustness checks.

[16] Panel (d) of Appendix Figure C-1 demonstrates that countries with multiple regime changes have comparatively attenuated trade effects, with the maximum treatment effect below 25% compared with over 75% for the single-regime change countries. Robust mean ATETs for single and multiple regime change countries are 0.19 ($t = 6.72$, $N = 4,806$) and 0.10 ($t = 4.76$, $N = 1,969$), respectively.

when we estimate our gravity model separately for the Cold War (1950-1989) and post-Cold War (1990-2015) periods. Recent work by Gokmen (2017) has demonstrated that before the collapse of the Soviet Bloc trade flows were strongly distorted by ideological affinity. Our results support this narrative: other than a positive effect for countries with 1-5 years in democracy, democratic regime change during the Cold War era had little discernible effect on exports. For the more recent period we again detect the inverted U-shape for countries with up to 15 years in democracy; having spent more time in democracy yields positive significant effects albeit below 20% (full sample).[17]

Robustness Our gravity equation includes dyadic measures for trade frictions: dummies for free trade agreements (FTAs) and common currency arrangements. These are prime objects of analysis in the current pooled structural gravity model paradigm.[18] However, in the context of a treatment effects analysis, these dummies could be seen as 'bad controls' (Angrist & Pischke 2008), in that democratic regime change may for instance result in an increased propensity to enter FTAs (though Baier & Bergstrand 2004, found no evidence for 'institutions' determining entry into FTAs). Therefore, in panel (b) of Appendix Figure C-1 we report results excluding these dummies/controls, with results next to identical to those in the baseline results.

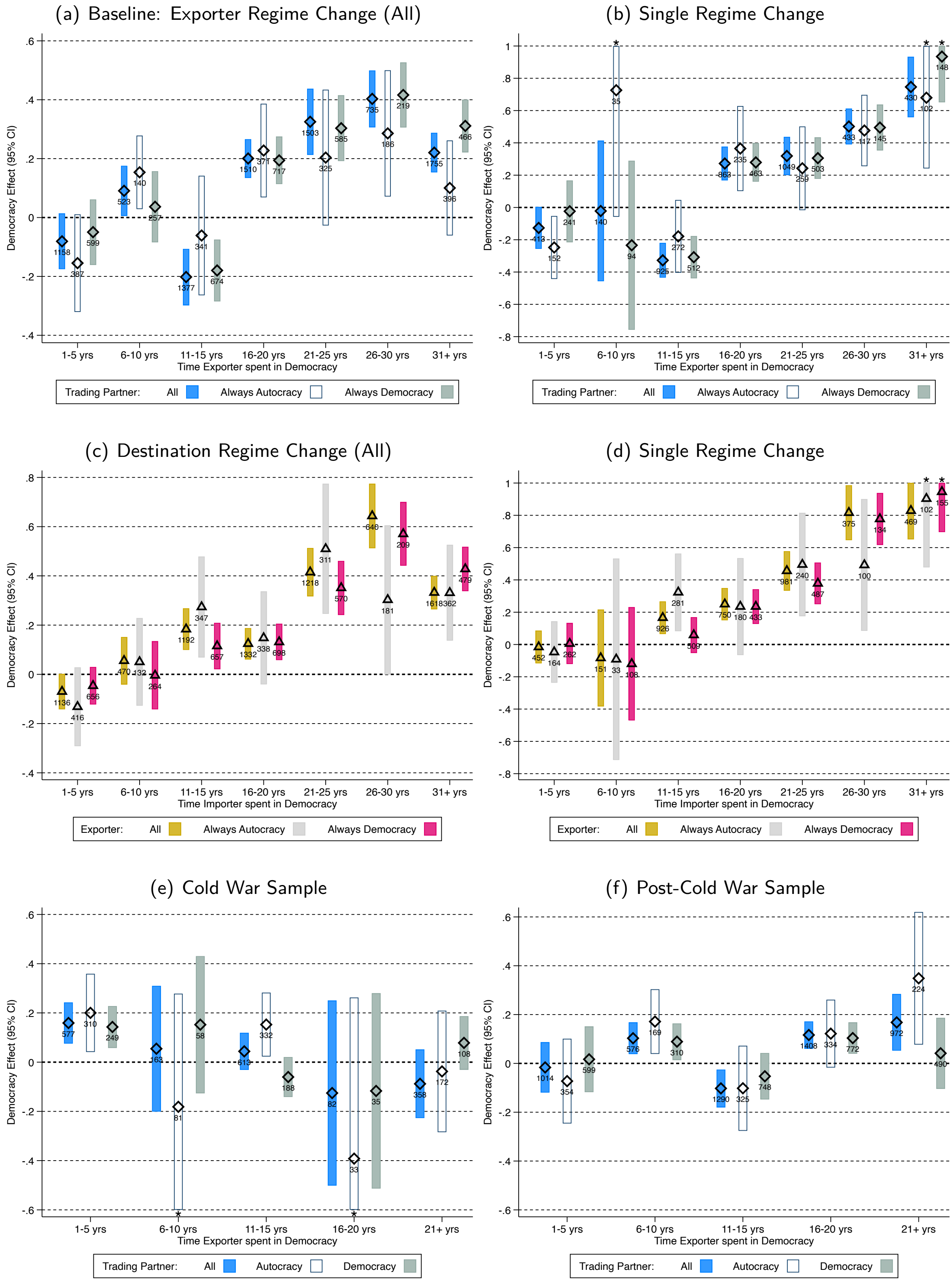
One feature of our implementation is the adjustment for treated countries which had zero export flows with specific destinations throughout the entire sample period. In Panel (e) of Appendix Figure C-1 we illustrate what happens to our results if we ignore this selection problem (excluding the null effect of democratic regime change) by omitting the final row of terms in equation (1) (except the error term). The patterns match those of our baseline results, but all magnitudes are inflated: after 25 (30) years the average treatment effect is over 60% (70%) as opposed to 30% (40%). Note that the Pseudo-Alpha test rejects this specification at the 1% level and we conclude that our zero-trade flow correction is crucially important for identification.

All of the above results are based on taking an arbitrary (mean) cut-off of the liberal democracy index to dichotomise regimes. As a robustness check, we still use the 1950-2014 index mean but add or subtract one-quarter or one-half of its standard deviation to define regime change. Appendix Figure C-2 displays more linear 'trajectories' for treatment effects over the years spent in democracy. If we use a more liberal (conservative) cut-off for regime change, effect sizes are typically smaller (larger). Notably, Pseudo-Alpha tests reject the two more liberal specifications ($p = 0.02$, respectively). For the most conservative democracy threshold in panel (e), the statistically significant

[17] There are at least two important caveats: first, the 1990-2015 period is much shorter and includes the Global Financial Crisis, thus naturally attenuating the potential for growth effects. Second, while the time series is much reduced, we still run the same demanding empirical specifications for this sub-period.

[18] In our results the mean coefficient on entering into an FTA has a very reasonable magnitude: around 0.15 ($t = 5.83$, $N = 1,375$) if we compute an outlier-robust mean in the sample of democratising exporters and around 0.13 ($t = 6.30$, $N = 2,299$) for all FTA estimates, suggesting 14-16% higher exports. The robust mean for common currency is -0.21 ($t = -1.92$, $N = 228$) among democratising exporters and -0.23 ($t = -2.40$, $N = 400$) for all estimates — most of these estimates (note the small sample sizes) are for regional African monetary unions.

Figure 1: Democratic Regime Change and Trade Flows — PPML-CCE-DID Results



Notes: These plots present ATET (using robust mean estimates) by time spent in democracy using hollow diamonds markers. These are not event plots: e.g. exporters with 21 years in democracy only appear in the ATET for 21-25 years. The bars indicate the 95% confidence interval and the underlying number of pairwise estimates is indicated. A value of 0.2 suggests that the switch from autocracy to democracy caused a 20% increase in exports. ∗ We curtailed the confidence interval for ease of presentation. Medians length in liberal democracy for all countries, destination autocracies and destination democracies are 20, 16, and 20 years, respectively. Panels (b) and (d) restrict the sample to countries which shifted from autocracy to democracy once (and did not revert). Panels (c) and (d) estimate the causal effect on exporter trade flows of *destination* country democratic regime change. Panels (e) and (f) are from separate gravity models for early and late time periods.

effects of democracy already manifest themselves for countries with 6-10 years in democracy, with maximum treatment effect over 60%. Estimates for sub-samples by destination regime status share the characteristics of the baseline results.

We also adopt the categorisation of Acemoglu et al. (2019) which is based on Polity IV and Freedom House definitions, in addition to detailed case analysis à la Papaioannou & Siourounis (2008). Results in panel (f) of Appendix Figure C-2 share many characteristics with our results for conservative thresholds using the V-Dem Liberal Democracy index (e.g. lacklustre/insignificant effects in the first 15 years, 'long-run' effect at 60%, higher variance of results for autocratic destinations).

4.2 Extensions

Average Trajectory of Treatment Effects Our results above relate the treatment effect to the time spent in democracy. The estimates for a country with, say, 23 years in democracy are *only* contained in the robust mean for '21-25 years' — we cannot make any statements about the trajectory of treatment effects for said country from the early years in democracy to the 23 years they eventually accumulated. In panel (f) of Appendix Figure C-1 we follow the strategy in Boese-Schlosser & Eberhardt (2024) and add dummies for the first fifteen years in democracy to our gravity regressions and present outlier-robust means for each of these first 15 years in democracy as well as for all years over 15. This mimics an event study analysis, at least for the initial periods, in the context of our heterogeneous gravity model.[19] The export effects of regime change are moderate and/or insignificant in the first half-dozen or so years (but *not* negative significant) and then rise to around 40% by 15 years in democracy. The average country with 16-62 years in democracy enjoys 22% higher exports. These results suggest that our earlier findings of an inverted U-pattern in the first 15 years are not representative of the trajectory of the *average* democratiser, but reflect the characteristics of the countries *with only few years in democracy*. Note that the two findings are not contradictory, they simply measure different things.

Building Blocks of Liberal Democracy Political scientists commonly adopt minimal definitions of democracy which build on Dahl's (1971) work and emphasise democracy as a "competitive struggle for the people's vote" (Schumpeter 1942, 26). They argue that the accountability of a regime can only come from the power of the electorate to extend or withdraw the mandate of the leader(s): "Democracy is a system in which parties lose elections" (Przeworski 1991, 10). Economists commonly put greater emphasis on the 'Northian' institutions (North 1981, North & Weingast 1989), a suite of civil rights (the 'rule of law') and constraints on the executive. These institutional building blocks ensure that the government is providing the right incentives and opportunities to private businesses,

[19] Note that the pairwise regressions now have twice as many parameters as those in our baseline model.

which entails the reduction of market frictions (e.g. financial development) and the facilitation of transactions more generally, most notably foreign trade (Besley 1995).

In our empirical analysis these building blocks are represented by the V-Dem proxies for 'electoral democracy' and the 'liberal component' — see Boese-Schlosser & Eberhardt (2023). We dichotomise these indices at their full sample means to construct treatment dummies for 'regime change'.[20]

Our analysis in panels (b) and (c) of Appendix Figure C-3 reveals that these different 'components' primarily speak to how countries with fewer than 10-15 years in democracy are affected: if we focus on electoral aspects, these countries have insignificant or marginally negative export effects, whereas if we emphasise the rule of law and executive constraints the negative export effects are significant (and sizeable). Treatment estimates for more mature democracies in contrast are very similar.

Systemic Patterns in Treatment Heterogeneity While the substantial causal effect of democracy on economic prosperity is well-established, we lack insights into effect heterogeneity across countries and potential drivers of this heterogeneity (Eberhardt 2022).

Table 1 reports robust mean treatment effects based on separate PPML-CCE-DID regressions where treatment and control samples are confined to identical characteristics in terms of a wide range of proxies for deep determinants of comparative development related to geography, legal origin, culture, and colonial history. The intuition is that we investigate the treatment effect of regime change in a country with specific structural characteristics (e.g. high historical disease environment) against a set of control countries with the same structural characteristics. This enables us to take the effect of structural characteristics on regime change, i.e. the notion that certain types of 'culture' or 'history' affect institutional quality in the long-run (e.g. Acemoglu et al. 2001), out of the equation and ask: how would democratic regime change in a country with good (or poor) deep determinants affect trade flows relative to a country with the same structural characteristics which remained autocratic?

We focus on the average treatment effect on the treated (ATET) per year spent in democracy, estimated from the gravity model results using an M-estimator (Rousseeuw & Leroy 2005):

$$\begin{aligned} \hat{\beta}_{ij}^{Dem} \quad = \quad & \theta_0 + \theta_1 \, \text{Years in Democracy}_i + \theta_2 \, \text{Years in Democracy}_i \times \text{Deep}_i \\ & + \theta_3 \, \text{ihs(Exports)}_{ij}^{Dem} + \theta_4 \, \text{ihs(Exports)}_{ij}^{Dem} \times \text{Deep}_i + \Psi' \text{Start Year}_{ij} + \varepsilon_{ij}, \end{aligned} \quad (2)$$

where $\hat{\beta}_{ij}^{Dem}$ is the treatment effect of democratic regime change for exporter i from the PPML-CCE-DID regression of trade flows from i to j for treatment and control samples with the same structural characteristic. Deep_i is a dummy for the deep determinant of development and takes the value 1 if country i is in the 'unfavourable deep determinant' sample and zero otherwise. $\text{ihs(Exports)}_{ij}^{Dem}$ is the value of exports from i to j in the year country i experienced regime change — since this can

[20] Regime change to electoral democracy and regime change to a 'high' liberal component.

be zero we adopt an inverse hyperbolic sine (ihs) transformation. This export term is also interacted with the 'unfavourable deep determinant' dummy.[21] Finally, Start Year$_{ij}$ is a set of dummies for the year the data for trade flows from i to j begins.

The regression in (2) is run separately by deep determinant and we report the annual change in exports per year in democracy in the favourable ($\hat{\theta}_1$) and unfavourable ($\hat{\theta}_1 + \hat{\theta}_2$) deep determinant samples, respectively, alongside a p-value indicating whether these estimates are statistically significantly different from each other (based on the t-ratio of $\hat{\theta}_2$).

In Table 1 Panel A for 1950-2014 almost all estimates of the trade effect of democratic regime change per additional year in democracy in columns (1) or (2) are statistically significant: we strike out any estimate if this is not the case, e.g. ~~1.1%~~ for being landlocked in column (2). Regardless of favourable or unfavourable geographic, legal, cultural, or historical characteristics, experiencing democratic regime change almost *unambiguously* leads to higher exports compared with countries with the same structural characteristics which remained autocratic. The magnitudes of the annualised effects are sizeable since the average country experiences over twenty years in democracy: democracy does cause economic prosperity, including through the channel of increased exports.[22]

Turning to the *differences* by deep determinant in column (3), we note that many specifications indicate statistically significant deviations in the export effect of democracy (p-values in the final column). Countries with unfavourable geography often have lower democracy effects (collectively -0.6%pa) than their peers with favourable characteristics, frequently statistically significantly so. The estimates by legal origin suggest that countries with unfavourable French legal origin (argued to be restrictive to the conduct of business and financial development) experience substantially *higher* trade growth following regime change.[23] Estimates by culture and colonial experience again suggest, by and large, that 'unfavourable' characteristics translate into a worse economic dividend of democratic regime change, though much less substantially so than in the case of geography (collectively -0.1%pa and -0.3%pa).

This investigation ignores the serious distortion of trade before the end of the Cold War, when ideology was a strong determinant of trade flows (Gokmen 2017). We re-run our analysis for the post-Cold War sample and present results in Panel B of the same table. The vast majority of estimates in columns (5) and (6) are positive and significant, but magnitudes are now substantially larger and patterns of differences have shifted somewhat: while results for geography and legal origin are similar to before and colonial experience has mixed results, we see a very strong differential effect

[21] Conditioning on trade flows in the regime change year anchors these estimates for the annual democracy trade effects and allows us to compare the treatment effect across samples of favourable and unfavourable deep determinants.

[22] Countries with favourable deep determinants on average have 22 years in democracy, 1.2 more than their peers. This gap differs by proxy, but all subsamples have on average either 16-20 or 21-25 years in democracy: the heterogeneous effects we find below cannot be attributed to countries with unfavourable deep determinants being largely stuck in the doldrums of the first 15 years in democracy.

[23] It should be noted that the suggestion of a detrimental effect of French legal origin (Porta et al. 2008) has been undermined by studies showing a reverse effect in historical data (Monnet & Velde 2021).

Table 1: Deep Determinants and the Democracy-Trade Effect

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		A: Full sample (1950-2014)				**B: Post-Cold War (1990-2014)**			
		Unfavourable Deep Determinant of Development							
Proxies		0	1	Diff	p(Diff)	0	1	Diff	p(Diff)
Geography									
Malaria Ecology	High	0.9%	1.4%	0.5%	0.07	0.8%	1.4%	0.5%	0.49
Malaria Risk	High	1.4%	1.2%	-0.2%	0.55	2.4%	2.9%	0.5%	0.49
Historical Disease Prevalence	High	1.8%	1.6%	-0.2%	0.48	2.6%	1.4%	-1.2%	0.14
Early Disease Environment (Auer)	High	~~-0.2%~~	0.6%	0.9%	0.00	3.5%	2.4%	-1.1%	0.12
UV Radiation	High	1.9%	0.2%	-1.6%	0.00	2.6%	1.0%	-1.6%	0.04
Absolute Latitude	Low	2.0%	0.7%	-1.3%	0.00	2.3%	~~0.0%~~	-2.3%	0.00
Landlocked	Yes	1.8%	~~1.1%~~	-0.7%	0.34	2.1%	~~0.8%~~	-1.3%	0.19
Few frost days per year	Yes	1.0%	~~-0.1%~~	-1.0%	0.00	3.2%	~~-0.2%~~	-3.4%	0.00
Some land in Tropical Zone	Yes	1.4%	1.3%	-0.2%	0.58	1.3%	3.3%	1.9%	0.01
No land in Temperate Zone	Yes	1.3%	1.0%	-0.3%	0.22	2.2%	1.5%	-0.7%	0.30
Unfavourable Geography	Yes	1.5%	0.9%	-0.6%	0.00	2.3%	1.4%	-0.6%	0.00
Legal Origin									
French Legal Origin	Yes	0.4%	1.4%	1.0%	0.00	~~-0.3%~~	1.9%	2.2%	0.00
Non-British Legal Origin	Yes	~~0.1%~~	1.9%	1.8%	0.00	~~0.8%~~	2.0%	1.1%	0.12
Unfavourable Legal Origin	Yes	0.3%	1.7%	1.4%	0.00	0.2%	2.0%	1.8%	0.00
Culture									
Collectivist	Yes	0.8%	1.7%	0.9%	0.00	5.7%	-1.0%	-6.7%	0.00
Blood type distance to UK	High	0.7%	1.1%	0.5%	0.09	1.5%	~~1.2%~~	-0.3%	0.66
Blood type distance to US	High	0.9%	1.0%	0.2%	0.58	1.8%	~~1.6%~~	-0.1%	0.84
Avg Common Language Index	Low	0.7%	~~0.2%~~	-0.5%	0.07	3.4%	0.8%	-2.6%	0.00
Avg Language Similarity	Low	1.4%	~~0.3%~~	-1.1%	0.00	1.9%	0.3%	-1.6%	0.02
Zero Euro descendants	Yes	1.7%	0.9%	-0.7%	0.01	3.9%	0.5%	-3.4%	0.00
No European settlers in 1900	Yes	0.8%	1.1%	0.3%	0.31	1.1%	~~1.3%~~	0.2%	0.80
Unfavourable Culture	Yes	1.0%	0.9%	-0.1%	0.16	2.4%	0.8%	-1.6%	0.00
History									
Colonial Experience	Yes	1.3%	0.9%	-0.4%	0.18	0.7%	2.0%	1.3%	0.15
Early Colonisation (<c19)	Yes	0.9%	~~0.2%~~	-0.6%	0.07	1.3%	2.5%	1.2%	0.33
Late Independence (>1959)	Yes	1.5%	1.1%	-0.4%	0.31	3.3%	~~0.9%~~	-2.4%	0.00
Unfavourable History	Yes	1.2%	0.9%	-0.3%	0.07	2.1%	1.7%	-0.5%	0.29

Notes: The table presents robust mean estimates of the export effect of democratic regime change per additional year in democracy (full sample, post-Cold War sample), columns (1) and (2), (5) and (6). These are derived from pooled models of the pairwise treatment effect regressed on (i) years in democracy, (ii) the start year of the pair sample, and (iii) the trade flow value (from exporter to importer) in the year the exporter experienced democratic regime change (transformed using the inverse hyperbolic sine). We distinguish robust mean estimates by proxies for deep determinants related to geography, legal origin, culture, and history. All of these proxies are dichotomous and coded so that a value of 1 indicates unfavourable deep determinants (high disease environment, French legal origin, etc), whereas 0 indicates favourable deep determinants. Columns (2) and (6) report the robust mean estimates for countries with unfavourable deep determinants, columns (1) and (5) for favourable — all of these estimates are statistically significantly different from zero, except those highlighted using 'strikethrough'. Columns (3) and (7) report the difference in means (unfavourable minus favourable mean estimate), and columns (4) and (8) whether this difference is statistically significant (p-value). We use darker shading for proxies which show statistical significance in the difference between unfavourable and favourable deep determinants in the democracy-trade effect. Columns (1) to (4) refer to the analysis of the full sample, (5) to (8) is for the post-Cold War sample (1990-2014).

of regime change by cultural characteristics (collectively -1.6%pa).[24] The end of the ideology-driven Cold War period has led to the emergence of cultural affinity as a major determinant of the trade effects of democratic regime change. In line with Gokmen's (2017, 449) findings we can conclude that "cultural determinants... [emerge] as a major impediment to international trade", even in the aftermath of democratic regime change.

5 Concluding Remarks

In this paper we introduced an empirical approach which seeks to satisfy the rigour of the structural gravity model and enable the analysis of monadic (country-specific) variables within a 'heterogeneous' gravity model. This opens up the opportunity to study a vast array of trade determinants which the current literature had to ignore in its adherence to the empirical straitjacket of the pooled gravity model. We employ this new machinery to investigate one of the most pertinent questions in development economics, namely that of the impact of institutional change on economic prosperity: does democracy cause greater trade flows? Our first set of results for the full sample from 1950 to 2014 indicate that the average country with three decades of democratic experience can expect high economic benefits in form of a 40% increase in exports. This effect is driven by increased exports to other democracies, whereas the magnitude of change in exports to autocracies is a lot more varied.

However, this finding masks substantial heterogeneities over time and across countries. If we limit the analysis to the Cold War era we see that regime change has virtually no economic effect. Causal effects of regime change after the Cold War are markedly lower than over the full sample period, at best 20% higher than in the counterfactual sample of autocracies. Economic benefits from regime change also follow distinct patterns of geographic, cultural, and historical 'bias'. In line with existing research, we find that these non-economic factors have come to the fore during the post-Cold War period. While regime change is unambiguously associated with higher export flows, countries which have unfavourable geography and comparatively distant culture have substantially lower effects than their peers with more favourable deep determiannts. Results for colonial history are mixed, whereas export effects are larger in countries with supposedly unfavourable French legal origin.

This is just a first attempt at highlighting the heterogeneous causal effects of institutional changes in the context of the empirical gravity model. Given the rich literature in political economy and political science, we expect the PPML-CCE-DID methodology to offer important insights in this regard in future research. The wider implications of a heterogeneous gravity model, allowing for the study of time-varying monadic variables, has of course significant potential to inform policy on a range of trade determinants which previously had to be ignored.

[24] Note that average length of treatment for countries with favourable (collectively 12.8 years) or unfavourable (collectively 14.5 years) deep determinants is uniformly in the 11-15 year bracket for most proxies.

References

Acemoglu, D., Johnson, S. & Robinson, J. A. (2001), 'The colonial origins of comparative development: An empirical investigation', American Economic Review **91**(5), 1369–1401.

Acemoglu, D., Johnson, S. & Robinson, J. A. (2005), 'The rise of europe: Atlantic trade, institutional change, and economic growth', American Economic Review **95**(3), 546–579.

Acemoglu, D., Naidu, S., Restrepo, P. & Robinson, J. A. (2019), 'Democracy does cause growth', Journal of Political Economy **127**(1), 47–100.

Allen, T., Arkolakis, C. & Takahashi, Y. (2020), 'Universal gravity', Journal of Political Economy **128**(2), 393–433.

Álvarez, I. C., Barbero, J., Rodríguez-Pose, A. & Zofío, J. L. (2018), 'Does institutional quality matter for trade? institutional conditions in a sectoral trade framework', World Development **103**, 72–87.

Anderson, J. E. & Van Wincoop, E. (2003), 'Gravity with gravitas: a solution to the border puzzle', American Economic Review **93**(1), 170–192.

Angrist, J. D. & Pischke, J.-S. (2008), Mostly harmless econometrics, Princeton University Press.

Arkolakis, C., Costinot, A. & Rodríguez-Clare, A. (2012), 'New trade models, same old gains?', American Economic Review **102**(1), 94–130.

Bai, J. (2009), 'Panel data models with interactive fixed effects', Econometrica **77**(4), 1229–1279.

Baier, S. L. & Bergstrand, J. H. (2004), 'Economic determinants of free trade agreements', Journal of International Economics **64**(1), 29–63.

Baier, S. L. & Bergstrand, J. H. (2007), 'Do free trade agreements actually increase members' international trade?', Journal of International Economics **71**(1), 72–95.

Baier, S. & Standaert, S. (2024), 'Gravity, globalization and time-varying heterogeneity', European Economic Review **163**, 104679.

Baum, M. A. & Lake, D. A. (2003), 'The political economy of growth: democracy and human capital', American Journal of Political Science **47**(2), 333–347.

Becker, B. (2019), Introducing coldat: The colonial dates dataset. socium, Sfb1342 working paper series, 02, Bremen: SOCIUM.

Besley, T. (1995), 'Property rights and investment incentives: Theory and evidence from ghana', Journal of Political Economy **103**(5), 903–937.

Beverelli, C., Keck, A., Larch, M. & Yotov, Y. V. (2024), 'Institutions, trade, and development: identifying the impact of country-specific characteristics on international trade', Oxford Economic Papers **76**(2), 469–494.

Boese-Schlosser, V. A. & Eberhardt, M. (2023), How does democracy cause growth?, Discussion Paper SP-V-2023-501, Berlin: WZB.

Boese-Schlosser, V. A. & Eberhardt, M. (2024), 'Democracy doesn't always happen overnight: regime change in stages and economic growth', Review of Economics and Statistics (forthcoming).

Boese, V. A. (2019), 'How (not) to measure democracy', International Area Studies Review **22**(2), 95–127.

Boneva, L. & Linton, O. (2017), 'A discrete-choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance', Journal of Applied Econometrics **32**(7), 1226–1243.

Cervellati, M. & Sunde, U. (2014), 'Civil conflict, democratization, and growth: Violent democratization as critical juncture', Scandinavian Journal of Economics **116**(2), 482–505.

Chan, M. K. & Kwok, S. S. (2022), 'The PCDID approach: difference-in-differences when trends are potentially unparallel and stochastic', Journal of Business & Economic Statistics **40**(3), 1216–1233.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L. et al. (2021), V-Dem Dataset v11, Varieties of Democracy Institute, Available at: https://doi.org/10.23696/vdemds21.

Costinot, A. & Rodríguez-Clare, A. (2014), Trade theory with numbers: Quantifying the consequences of globalization, in 'Handbook of International Economics', Vol. 4, Elsevier, pp. 197–261.

Dahl, R. A. (1971), Polyarchy: Participation and opposition, Yale University Press, New Haven and London.

Dutt, P. & Traca, D. (2010), 'Corruption and bilateral trade flows: extortion or evasion?', Review of Economics and Statistics **92**(4), 843–860.

Eaton, J. & Kortum, S. (2002), 'Technology, geography, and trade', Econometrica **70**(5), 1741–1779.

Eberhardt, M. (2022), 'Democracy, growth, heterogeneity, and robustness', European Economic Review **147**(104173).

Eberhardt, M., Helmers, C. & Strauss, H. (2013), 'Do spillovers matter when estimating private returns to R&D?', Review of Economics and Statistics **95**(2), 436–448.

Eberhardt, M. & Presbitero, A. F. (2015), 'Public debt and growth: Heterogeneity and non-linearity', Journal of International Economics **97**(1), 45–58.

Fally, T. (2015), 'Structural gravity and fixed effects', Journal of International Economics **97**(1), 76–85.

Feyrer, J. (2019), 'Trade and income—exploiting time series in geography', American Economic Journal: Applied Economics **11**(4), 1–35.

Fouquin, M. & Hugot, J. (2016), Two centuries of bilateral trade and gravity data: 1827-2014, CEPII Working Paper 2016-14.

Francois, J. & Manchin, M. (2013), 'Institutions, infrastructure, and trade', World Development **46**, 165–175.

Gerring, J., Bond, P., Barndt, W. T. & Moreno, C. (2005), 'Democracy and economic growth: A historical perspective', World Politics **57**(3), 323–364.

Gobillon, L. & Magnac, T. (2016), 'Regional policy evaluation: Interactive fixed effects & synthetic controls', Review of Economics and Statistics **98**(3), 535–51.

Gokmen, G. (2017), 'Clash of civilizations and the impact of cultural differences on trade', Journal of Development Economics **127**, 449–458.

Gorodnichenko, Y. & Roland, G. (2017), 'Culture, institutions, and the wealth of nations', Review of Economics and Statistics **99**(3), 402–416.

Head, K. & Mayer, T. (2014), Gravity equations: Workhorse, toolkit, and cookbook, in 'Handbook of International Economics', Vol. 4, Elsevier, pp. 131–195.

Head, K. & Ries, J. (2008), 'FDI as an outcome of the market for corporate control: Theory and evidence', Journal of International Economics **74**(1), 2–20.

Heid, B., Larch, M. & Yotov, Y. V. (2021), 'Estimating the effects of non-discriminatory trade policies within structural gravity models', Canadian Journal of Economics/Revue canadienne d'économique **54**(1), 376–409.

Huntington, S. P. (1991), 'Democracy's third wave', Journal of Democracy **2**(2), 12–34.

Imam, P. A. & Temple, J. R. (2024), 'Political institutions and output collapses', European Journal of Political Economy **85**, 102573.

Knutsen, C. H. (2012), 'Democracy and economic growth: A survey of arguments and results', International Area Studies Review **15**(4), 393–415.

Milner, H. V. & Kubota, K. (2005), 'Why the move to free trade? democracy and trade policy in the developing countries', International Organization **59**(1), 107–143.

Monnet, E. & Velde, F. R. (2021), Money, banking, and old-school historical economics, in 'The Handbook of Historical Economics', Elsevier, pp. 335–364.

Murray, D. R. & Schaller, M. (2010), 'Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture', Journal of Cross-Cultural Psychology **41**(1), 99–108.

North, D. C. (1981), Structure and change in economic history, New York: Norton.

North, D. C. & Weingast, B. R. (1989), 'Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century england', Journal of Economic History **49**(4), 803–832.

Nunn, N. & Puga, D. (2012), 'Ruggedness: The blessing of bad geography in africa', Review of Economics and Statistics **94**(1), 20–36.

Nunn, N. & Trefler, D. (2014), 'Domestic institutions as a source of comparative advantage', Handbook of International Economics **4**, 263–315.

O'Rourke, K. H. & Taylor, A. M. (2006), Democracy and protectionism, NBER Working Paper Series 12250, National Bureau of Economic Research.

Papaioannou, E. & Siourounis, G. (2008), 'Democratisation and growth', Economic Journal **118**(532), 1520–1551.

Persson, T. & Tabellini, G. (2009), 'Democratic capital: The nexus of political and economic change', American Economic Journal: Macroeconomics **1**(2), 88–126.

Pesaran, M. H. (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', Econometrica **74**(4), 967–1012.

Pesaran, M. H. & Smith, R. (1995), 'Estimating long-run relationships from dynamic heterogeneous panels', Journal of Econometrics **68**(1), 79–113.

Porta, R. L., Lopez-de Silanes, F. & Shleifer, A. (2008), 'The economic consequences of legal origins', Journal of Economic Literature **46**(2), 285–332.

Przeworski, A. (1991), Democracy and the market: Political and economic reforms in Eastern Europe and Latin America, Cambridge University Press.

Przeworski, A., Alvarez, M., Cheibub, J. A. & Limongi, F. (2000), Democracy and development: political institutions and well-being in the world, 1950-1990, Vol. 3, Cambridge University Press, Cambridge, UK.

Puga, D. & Trefler, D. (2014), 'International trade and institutional change: Medieval venice's response to globalization', Quarterly Journal of Economics **129**(2), 753–821.

Rousseeuw, P. J. & Leroy, A. M. (2005), Robust regression and outlier detection, John Wiley & Sons.

Santos Silva, J. & Tenreyro, S. (2006), 'The log of gravity', Review of Economics and Statistics **88**(4), 641–658.

Schumpeter, J. A. (1942), Capitalism, socialism and democracy, Harper & Brothers: London and New York.

Spolaore, E. & Wacziarg, R. (2013), 'How deep are the roots of economic development?', Journal of Economic Literature **51**(2), 325–369.

Swamy, P. A. (1970), 'Efficient inference in a random coefficient regression model', Econometrica pp. 311–323.

Xu, Y. (2017), 'Generalized synthetic control method: Causal inference with interactive fixed effects models', Political Analysis **25**(1), 57–76.

Yotov, Y. V. (2024), 'The evolution of structural gravity: The workhorse model of trade', Contemporary Economic Policy **42**(4), 578–603.

Yotov, Y. V., Piermartini, R., Larch, M. et al. (2016), An advanced guide to trade policy analysis: The structural gravity model, Geneva: World Trade Organization.

Yu, M. (2010), 'Trade, democracy, and the gravity equation', Journal of Development Economics **91**(2), 289–300.

Online Appendix – Not Intended for Publication

A Data, Sample Makeup, and Descriptives

A.1 Detailed sources of proxies for deep determinants

For the disease environment aspect, we adopt two datasets related to malaria: the percentage share of the population at risk of malaria in 1965 from McCord et al (2017) and malaria ecology, a measure of the latent risk of malaria, from Kiszewski et al (2004). We also use historical data collated by Murray and Schaller (2010) on the prevalence of seven parasitic and infectious diseases, and data from Auer (2013) on early (18th century) disease environment, extended in Ertan et al (2016). For climate-related aspects of geography, we use data collated by Nunn and Puga (2012) on average latitude and land area in the tropical climate zone to construct dummies for low absolute latitude (the region between the Tropic of Cancer and the Tropic of Capricorn) and 'some land in the tropics'. We further construct a dummy for no land in the temperate climate zone using information from Spolaore and Wacziarg (2013), from where we also collect information on whether countries are landlocked (unrelated to climate). From Andersen et al (2016) we take data on UV radiation intensity and information on frost days per year (originally constructed by Masters and McMillan, 2001).

Information on French legal origin is taken from La Porta et al (2008), from which we also construct a dummy for all countries which do not have British legal origin.

Proxies for culture relate to aspects of Hofstede's concept of 'individualism' (good for innovation and growth) versus 'collectivism' (less good), and related proxies with broader country coverage from Gorodnichenko and Roland (2017): to blood type distance to the UK and US (the two most individualistic countries in the world). We further study the ability to communicate, proxied by language similarity, which we construct as country averages from dyadic data by Gurevich et al (2021). Our proxies capture the probability that two individuals picked at random from each pair of countries speak the same native language (cnl) and a population-weighted proximity measure based on 'linguistic trees' which categorise languages (lp).[25] Following the ideas of Max Weber, dummies for 'no descendants from European settlers' today and 'no European settlers' at the turn of the 20th century based on data collated by Spolaore and Wacziarg (2013) and Gorodnichenko and Roland (2017) are our further proxies for culture.

For aspects of colonial history we use the COLDAT database (Becker, 2019) from which we construct dummies for colonial experience, early colonisation (before the 19th century), and late independence (after the 1950s). The rationale of the latter two is to gauge 'long exposure' to colonial domination.

A.2 Sample makeup

The baseline specification (liberal democracy with controls) has 8,568 estimates for regime change. In Appendix Table D-1 panel column (1) the robust mean ATET of 0.137 is computed using an M-estimator which weights each estimate (0 to 1). In our graphical presentation in the main text those estimates with zero weight are not counted/included, our total here is 7,732 pairwise estimates for 84 exporters (and up to 164 destinations).

61 countries (73%) had a single regime change from democracy to autocracy of countries), 12 had two, 6 three and 2 four — 3 countries only reverted from autocracy to democracy. Mean (median) time spent in democracy is 23 (22) years. Appendix Table A-1 provides further details.

[25] The DICL data by Gurevich et al (2021) are dyadic. However, to maintain the logic of our analysis by deep determinant (typically monadic), we are interested in and employ exporter i's *average* ability to communicate with all other countries, not their ability with specific importer j.

Table A-1: Sample Makeup: Exporters

	ISO	Exporter	Start	Pairs	Demo	Reversal	Years
1	ALB	Albania	1950	54	1	0	23
2	ARG	Argentina	1950	124	4	3	38
3	ARM	Armenia	1992	43	0	1	3
4	BEN	Benin	1959	64	1	0	24
5	BFA	Burkina Faso	1954	81	1	0	16
6	BGR	Bulgaria	1950	102	1	0	25
7	BIH	Bosnia & Herzegovina	1994	53	1	0	18
8	BLR	Belarus	1992	111	0	1	5
9	BOL	Bolivia	1950	64	1	0	30
10	BRA	Brazil	1950	132	1	0	29
11	BTN	Bhutan	1980	34	1	0	7
12	CHL	Chile	1950	95	2	2	46
13	CIV	Cote d'Ivoire	1960	134	1	0	13
14	COG	Republic of Congo	1950	62	2	2	2
15	COL	Colombia	1950	97	2	1	35
16	COM	Comores	1980	44	1	0	1
17	CPV	Cabo Verde	1975	24	1	0	24
18	CUB	Cuba	1950	36	0	1	1
19	CYP	Cyprus	1960	34	1	0	54
20	DOM	Dominican Republic	1950	90	2	1	22
21	ECU	Ecuador	1950	121	1	1	33
22	ESP	Spain	1950	129	1	0	37
23	FJI	Fiji	1960	80	3	3	32
24	GEO	Georgia	1992	79	1	0	11
25	GHA	Ghana	1950	101	4	3	27
26	GMB	The Gambia	1965	49	1	1	27
27	GRC	Greece	1950	113	2	1	51
28	GTM	Guatemala	1950	105	1	0	16
29	GUY	Guyana	1960	60	1	0	22
30	HKG	Hong Kong	1960	137	1	0	23
31	HND	Honduras	1950	114	1	1	14
32	HRV	Croatia	1995	114	1	0	15
33	HUN	Hungary	1950	121	1	0	25
34	IDN	Indonesia	1950	159	2	1	18
35	IND	India	1950	125	2	1	61
36	JAM	Jamaica	1950	15	1	0	62
37	KEN	Kenya	1955	152	2	1	10
38	KGZ	Kirgistan	1992	68	1	0	4
39	KOR	South Korea	1950	130	1	0	27
40	LBN	Lebanon	1950	150	1	0	4
41	LBR	Liberia	1950	87	1	0	9
42	LBY	Libya	1960	70	1	1	1
43	LKA	Sri Lanka	1950	150	2	3	45
44	LSO	Eswatini	1960	52	2	1	15
45	MDG	Madagascar	1950	119	1	1	3
46	MDV	Maldives	1980	58	1	1	4
47	MEX	Mexico	1950	120	1	0	18
48	MKD	North Macedonia	1993	83	1	1	18
49	MLI	Mali	1960	85	2	1	21
50	MNG	Mongolia	1950	36	1	0	24

(Continued Overleaf)

Table A-1: Sample Makeup: Exporters (continued)

	ISO	Exporter	Start	Pairs	Demo	Reversal	Years
51	MOZ	Mozambique	1975	84	1	0	20
52	MWI	Malawi	1954	88	1	0	20
53	NAM	Namibia	1980	21	1	0	25
54	NER	Niger	1960	94	3	2	16
55	NGA	Nigeria	1950	104	1	0	15
56	NIC	Nicaragua	1950	96	1	1	17
57	NPL	Nepal	1952	99	3	2	15
58	PAN	Panama	1950	83	1	0	24
59	PER	Peru	1950	130	3	2	29
60	PHL	Philippines	1950	120	1	0	27
61	PNG	Papua New Guinea	1960	32	1	0	43
62	POL	Poland	1950	127	1	0	25
63	PRT	Portugal	1950	119	1	0	39
64	PRY	Paraguay	1950	74	1	0	23
65	RUS	Russia	1992	129	1	2	2
66	SEN	Senegal	1959	69	1	0	37
67	SGP	Singapore	1950	152	1	0	15
68	SLB	Solomon Islands	1967	38	2	1	35
69	SLE	Sierra Leone	1960	79	1	0	12
70	SLV	El Salvador	1950	95	1	0	16
71	STP	Sao Tome & Principe	1975	18	1	0	16
72	SUR	Suriname	1960	46	1	1	47
73	SYC	Sychelles	1960	71	1	0	12
74	THA	Thailand	1950	162	3	3	18
75	TTO	Trinidad & Tobago	1951	44	1	0	59
76	TUN	Tunisia	1960	130	1	0	3
77	TUR	Turkey	1950	164	3	3	39
78	TWN	Taiwan	1951	146	1	0	14
79	TZA	Tanzania	1955	93	1	0	23
80	UKR	Ukraine	1992	154	1	2	10
81	URY	Uruguay	1950	95	1	1	53
82	VEN	Venezuela	1950	103	1	1	42
83	ZAF	South Africa	1950	126	1	0	20
84	ZMB	Zambia	1950	87	1	0	23
	Total			7732			

Notes: We present details of the 84 exporters which experienced regime change during 1950-2015 based on our baseline Liberal Democracy definition. 'Start' indicates the year in which the country enters the dataset, 'Pairs' the number of trade partners (with positive trade flows in at least one year), 'Demo' the number of democratic regime changes, 'Reversal' the number of autocratic reversals, and 'Years' the total number of years in democracy. We only report these statistics for countries which did not get assigned a zero weight by our M-estimator.

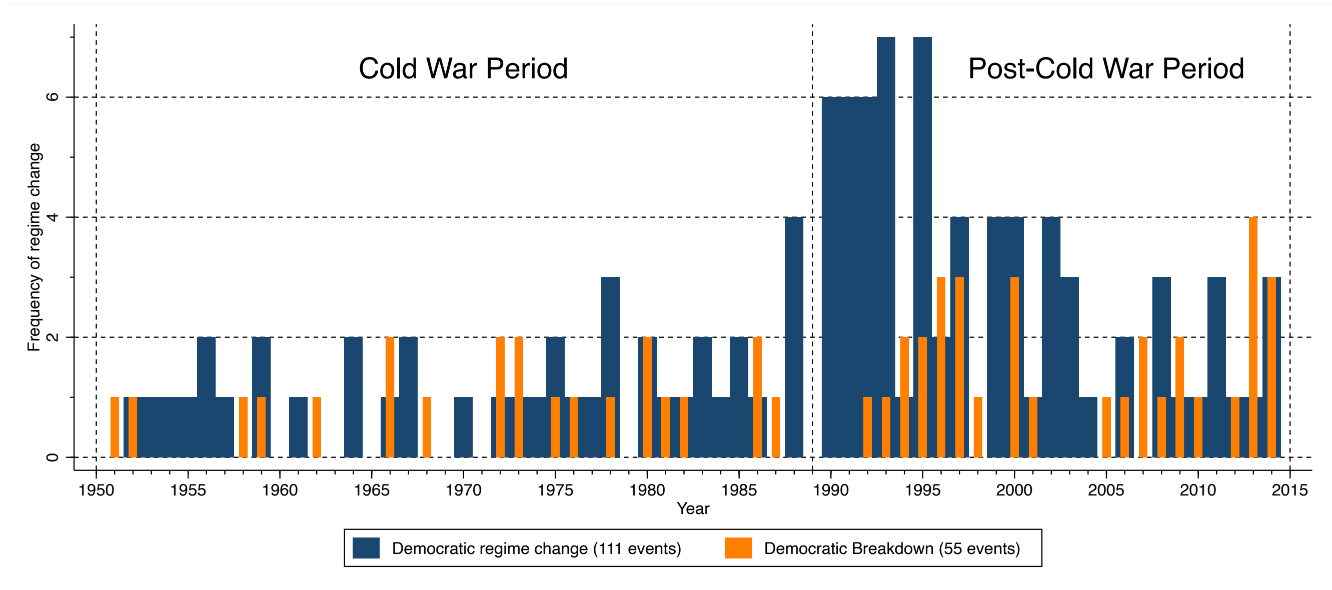
A.3 Descriptives

Panel (a) of Appendix Figure A-1 provides the distribution of regime change events (111 democratisations, 55 autocratisations) in the exporter countries of our sample. We distinguish the Cold War and Post-Cold War periods and highlight the 'Third Wave of Democratisation' (Huntington 1991) in the years following 1989 (33 democratic regime change events during 1990-95).

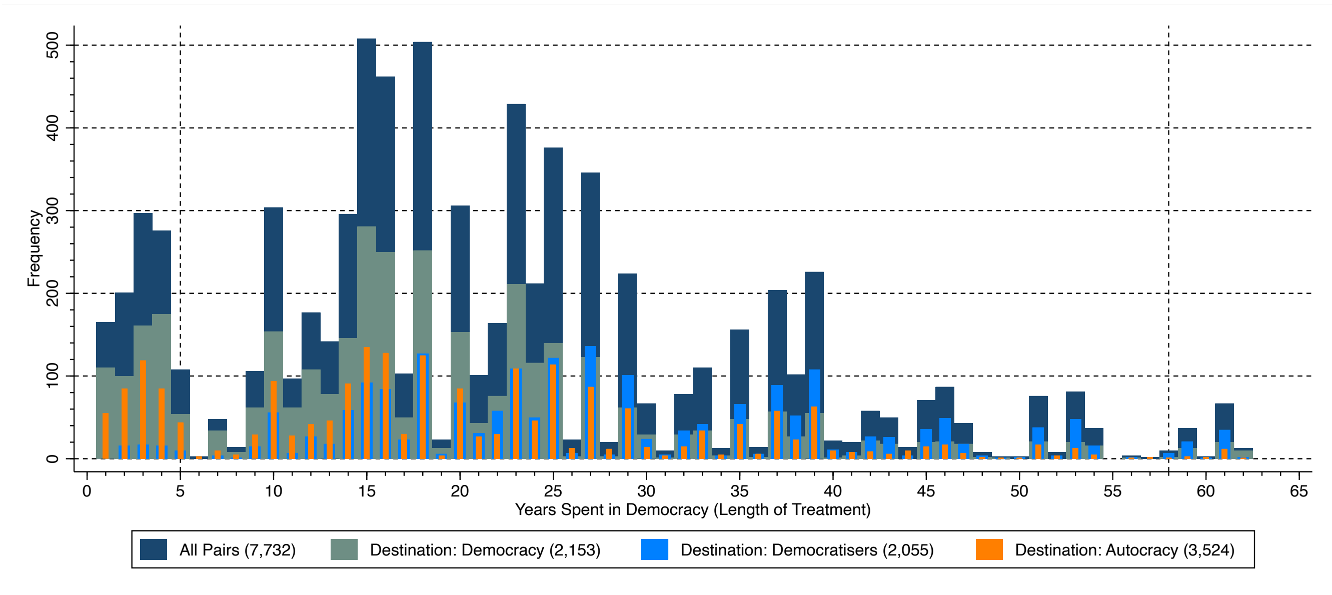
Panel (b) indicates the distribution of years spent in democracy (using the dyadic data) for four different samples, defined by the destination: (i) all pairs, (ii) pairs where the importer is a democracy throughout the sample period, (iii) pairs where the importer experiences regime change during the sample period, and (iv) pairs where the importer is autocratic throughout. In this and all results plots by construction the country experiencing regime change is always the exporter i, and on the x-axis we can read off the total number of years the exporter country has spent in democracy. With exception of the democratiser sub-sample,[26] these distributions suggest that the thousands of democracy estimates making up the results we present below are relatively uniformly distributed across these different periods in democracy,[27].

Figure A-1: Democratic Regime Change — Years (countries) and Length of Treatment (dyadic)

(a) Year of Regime Change (Liberal Democracy)



(b) Distribution — Years Spent in Democracy (Liberal Democracy)



Notes: In panel (a) we present the frequency of regime change by year for the 84 (exporter) countries which experienced regime change. A total of 111 democratic regime changes and 55 democratic breakdowns occurred during the 1950-2014 sample period. Panel (b) uses the dyadic data for these 84 countries (7,732 pairs) and indicates the frequency of total number of years spent in democracy on the x-axis.

[26]Here the youngest democracies only make up 7% of the sample, 24% of estimates are for countries with 11-20 years of experience, and 35% each for 21-30 and over 30 years, respectively.

[27]Around one-quarter of estimates are for exporters with up to 10 years experience of democracy, 30-35% for those with 11 to 20 years, 24-28% for those with 21 to 30 years and 13-20% for those with more than 30 years in democracy.

A.4 Descriptives for the Deep Determinants Analysis

Table A-2: Treatment Years by Deep Determinant (Full Sample: 1950-2015)

Deep Determinant		Favourable				Unfavourable				
		N	Mean	Med	SD	N	Mean	Med	SD	Dev'n
Geography										
Malaria Ecology	High	3609	23.0	22	16.4	4957	20.0	18	9.9	-3.0
Malaria Risk	High	4149	22.1	22	14.6	4422	20.4	18	11.5	-1.8
Historical Disease Prevalence	High	3348	20.4	20	12.2	5229	21.8	18	13.6	1.4
Early Disease Environment (Auer)	High	3524	21.3	20	12.2	5072	21.2	18	13.7	-0.1
UV Radiation	High	3009	21.0	23	15.0	5580	21.3	18	12.0	0.3
Absolute Latitude	Low	3516	20.6	18	14.4	5049	21.8	20	12.1	1.2
Landlocked	Yes	7158	22.2	20	13.8	1419	16.6	18	7.6	-5.5
Few frost days per year	Yes	4207	19.9	18	13.9	4367	22.6	21	12.2	2.7
Some land in Tropical Zone	Yes	3291	21.2	23	14.8	5286	21.3	18	11.9	0.1
No land in Temperate Zone	Yes	4260	21.0	20	13.4	4305	21.5	18	12.9	0.4
Legal Origin										
French Legal Origin	Yes	3684	21.0	20	12.6	4886	21.5	19	13.5	0.5
Non-British Legal Origin	Yes	2503	22.6	20	14.0	6073	20.7	18	12.7	-1.8
Culture										
Collectivist	Yes	1904	28.1	29	17.4	3985	22.0	20	9.9	-6.0
Blood type distance to UK	High	4094	21.9	20	13.9	4485	20.6	18	12.4	-1.3
Blood type distance to US	High	3886	22.4	23	14.2	4701	20.2	18	12.1	-2.2
Avg Common Language Index	Low	4546	24.1	23	13.4	4031	18.1	16	12.0	-5.9
Avg Language Similarity	Low	4276	22.4	20	14.3	4299	20.2	18	11.7	-2.2
Zero Euro descendants	Yes	5519	22.3	22	13.2	3043	19.3	16	12.8	-3.0
No European settlers in 1900	Yes	5586	22.7	22	13.4	3009	18.6	16	12.2	-4.0
Colonial History										
Colonial Experience	Yes	2543	20.0	18	12.6	5776	21.9	20	13.6	1.9
Early Colonisation (<c19)	Yes	2916	16.7	15	13.0	3100	26.6	24	11.8	9.8
Late Independence (>1959)	Yes	3365	24.3	23	13.8	2663	18.6	16	11.8	-5.7

Notes: We present distributional statistics for the deep determinant analysis. N is the number of pair estimates for regime change, Mean indicates the average length of treatment (total years in democracy), Med the median, SD the standard deviation. Dev'n shows the deviation between average unfavourable and favourable length of treatment.

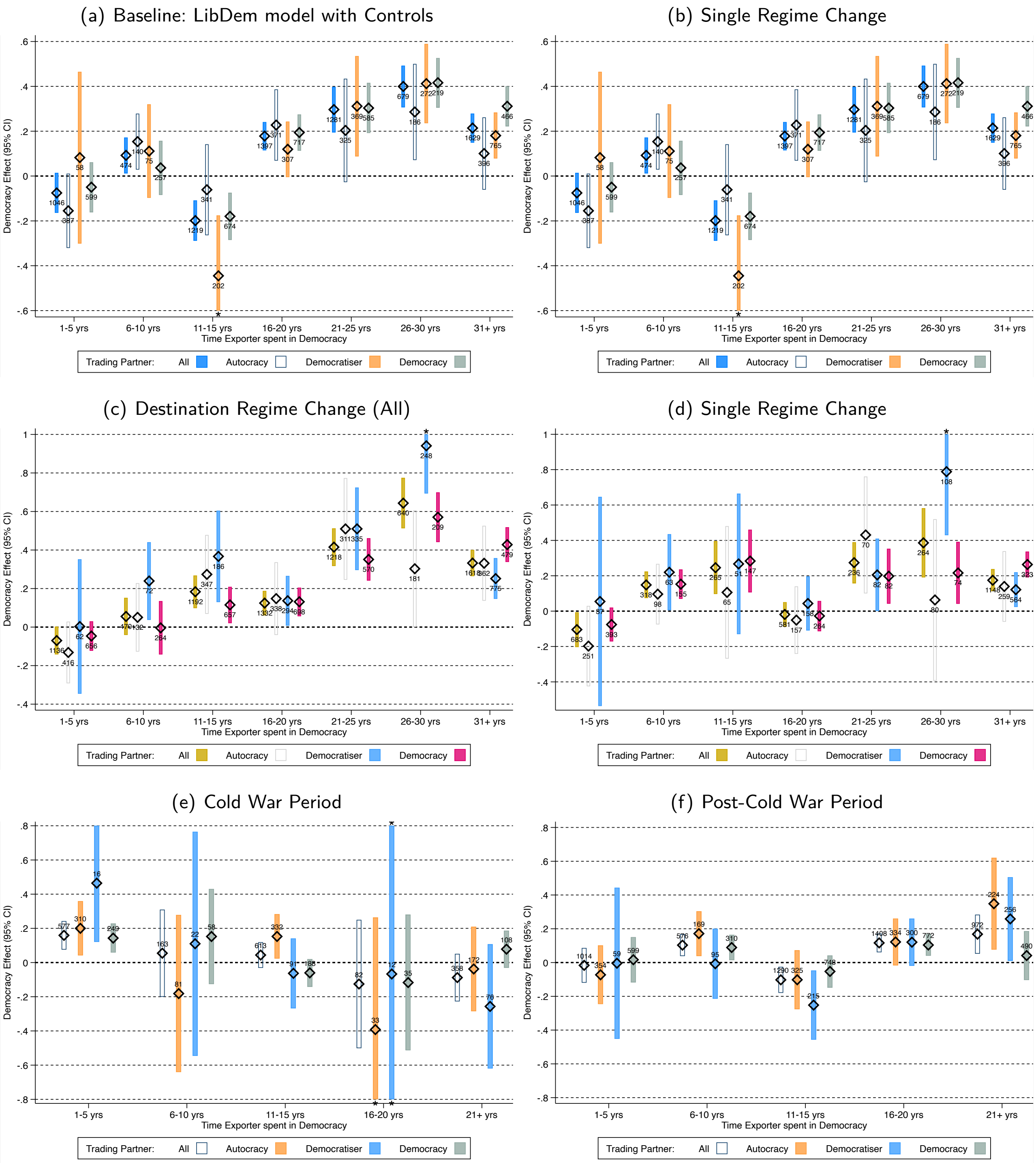
Table A-3: Treatment Years by Deep Determinant (Post-Cold War: 1990-2015)

Deep Determinant		Favourable				Unfavourable				
		N	Mean	Med	SD	N	Mean	Med	SD	Dev'n
Geography										
Malaria Ecology	High	2062	10.5	10	7.0	3823	15.5	16	5.8	5.1
Malaria Risk	High	2558	12.3	12	7.1	3334	14.9	16	6.1	2.7
Historical Disease Prevalence	High	2108	13.1	14	6.3	3780	14.2	15	6.9	1.1
Early Disease Environment (Auer)	High	2504	14.6	15	6.5	3407	13.2	15	6.8	-1.4
UV Radiation	High	1693	10.3	9	7.9	4237	15.1	16	5.6	4.8
Absolute Latitude	Low	2159	11.4	11	7.3	3736	15.2	16	5.9	3.7
Landlocked	Yes	4820	13.6	15	6.7	1085	14.5	16	6.8	0.9
Few frost days per year	Yes	2804	12.1	13	7.3	3095	15.3	16	5.7	3.3
Some land in Tropical Zone	Yes	1913	11.2	10	7.9	3971	15.0	16	5.6	3.8
No land in Temperate Zone	Yes	2683	13.0	14	7.5	3202	14.5	16	5.9	1.5
Legal Origin										
French Legal Origin	Yes	2586	14.6	15	5.7	3298	13.2	16	7.3	-1.4
Non-British Legal Origin	Yes	1979	15.2	15	5.7	3922	13.0	15	7.0	-2.2
Culture										
Collectivist	Yes	840	10.4	7	8.6	2821	15.9	16	4.2	5.5
Blood type distance to UK	High	2496	12.6	12	7.1	3402	14.7	16	6.3	2.1
Blood type distance to US	High	2192	11.9	11	7.1	3714	14.9	16	6.2	3.0
Avg Common Language Index	Low	2720	14.9	16	6.4	3144	12.8	14	6.8	-2.1
Avg Language Similarity	Low	2607	12.8	15	7.0	3253	14.6	15	6.3	1.8
Zero Euro descendants	Yes	3435	14.0	15	6.9	2461	13.5	14	6.5	-0.5
No European settlers in 1900	Yes	3535	13.8	15	6.8	2359	13.7	15	6.5	-0.1
Colonial History										
Colonial Experience	Yes	1579	12.3	14	6.7	4078	13.9	15	6.6	1.6
Early Colonisation (<c19)	Yes	2458	12.4	13	7.4	1837	16.8	16	4.2	4.4
Late Independence (>1959)	Yes	2174	14.7	16	6.9	2136	13.9	15	6.4	-0.8

Notes: We present distributional statistics for the deep determinant analysis limited to the post-Cold War period (1990-2015). N is the number of pair estimates for regime change, Mean indicates the average length of treatment (total years in democracy), Med the median, SD the standard deviation. Dev'n shows the deviation between average unfavourable and favourable length of treatment.

B Full Results

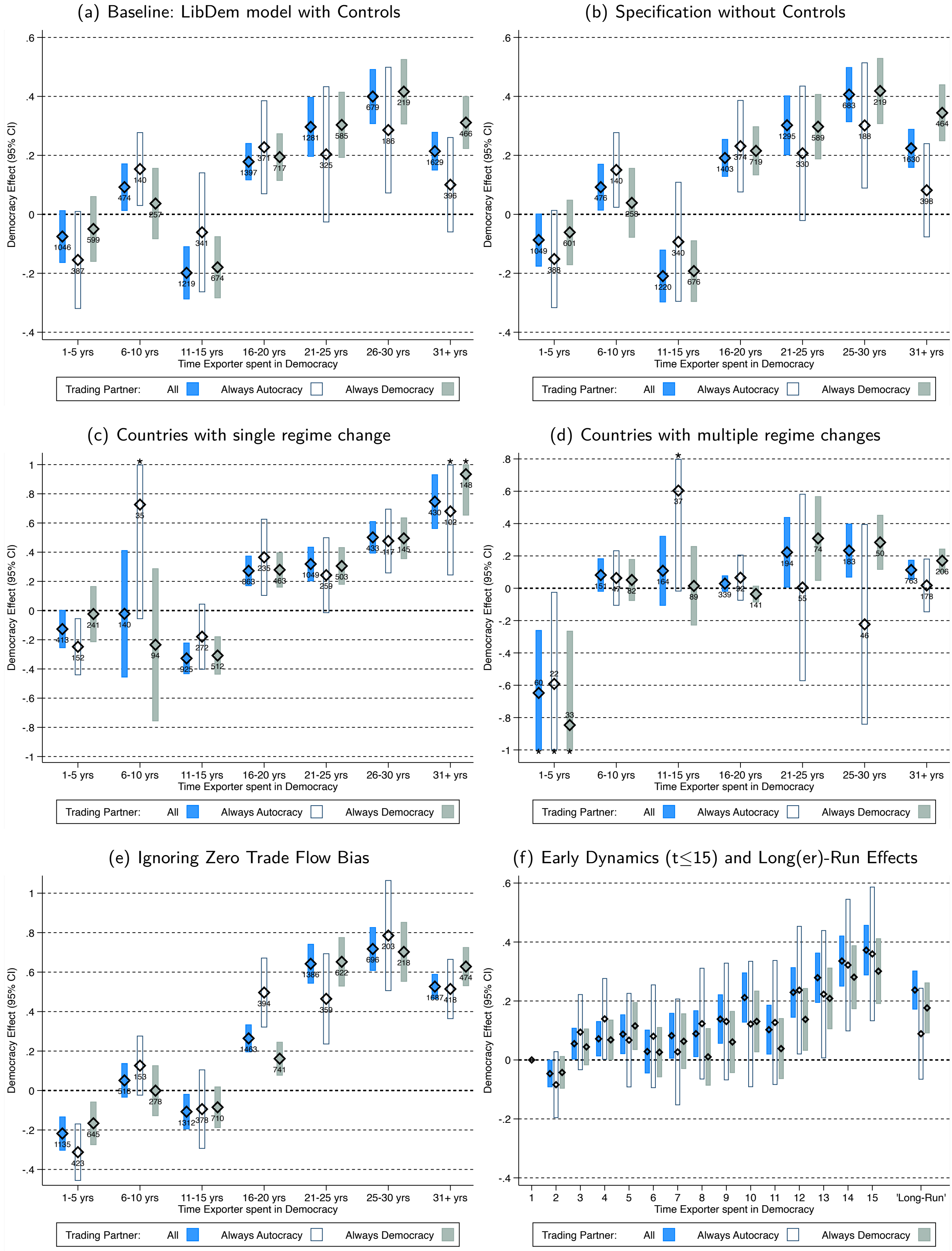
Figure B-1: Democratic Regime Change and Trade Flows — Including Democratising Destinations



Notes: These plots present robust mean estimates for time spent in democracy (5-year averages). We distinguish trade partners which are always democracies, always autocracies, or (unique to these plots) democratised themselves during the sample period. The bars are the 95% confidence intervals, the markers the outlier-robust means. We indicates the number of country pairs included in each estimate.

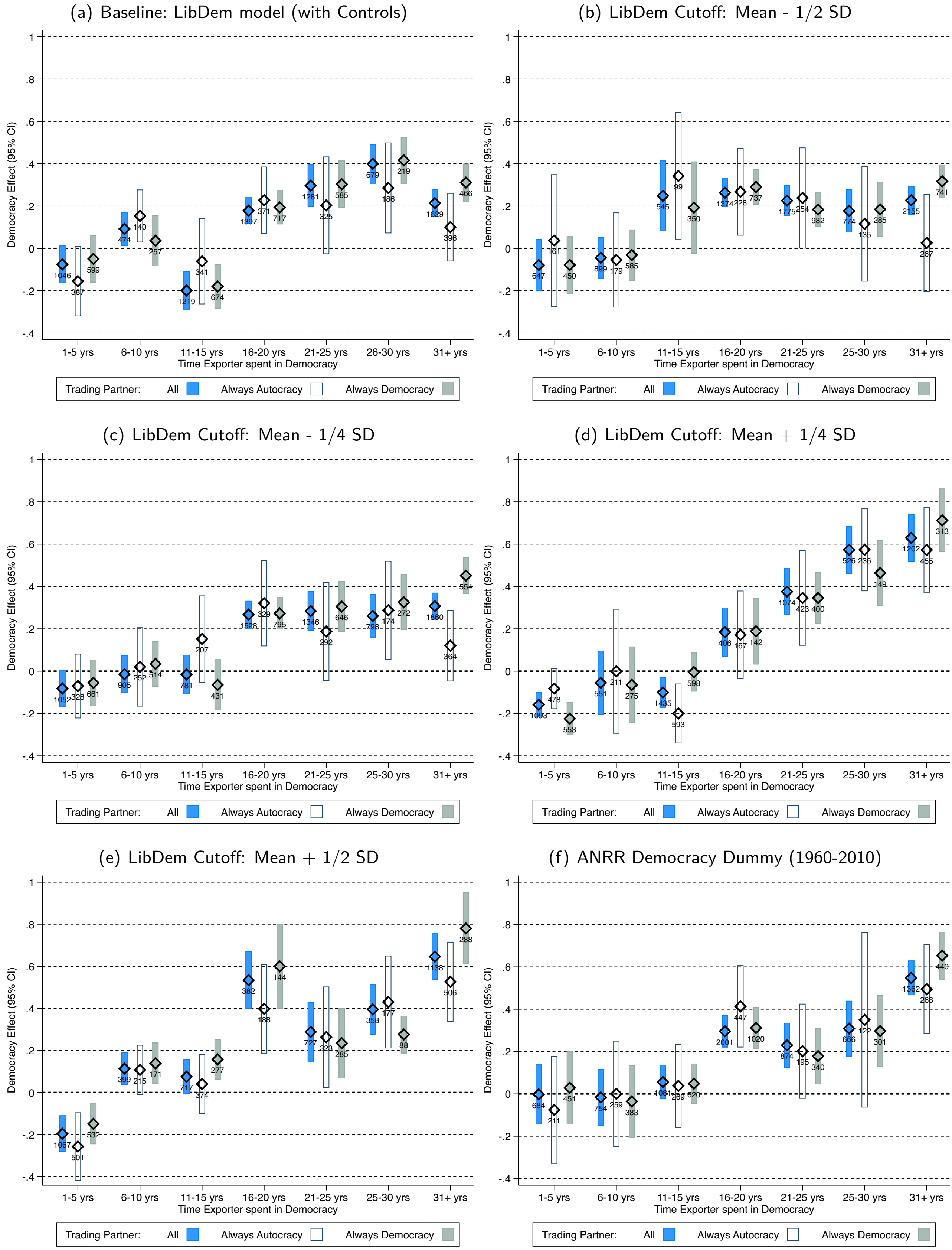
C Results: Robustness Checks

Figure C-1: Democratic Regime Change and Trade Flows — Various Specifications



Notes: These plots present robust mean estimates for time spent in democracy (5-year averages except in (f)). We distinguish trade partners which are always democracies and always autocracies. The bars are the 95% confidence intervals, the markers the outlier-robust means. We indicates the number of country pairs included in each estimate. Total estimates (full sample excluding zero-weight observations) amount to (a) 7,732 (b) 7,763, (c) 5,308, (d) 1,862, (e) 8,202, (f) 5,821.

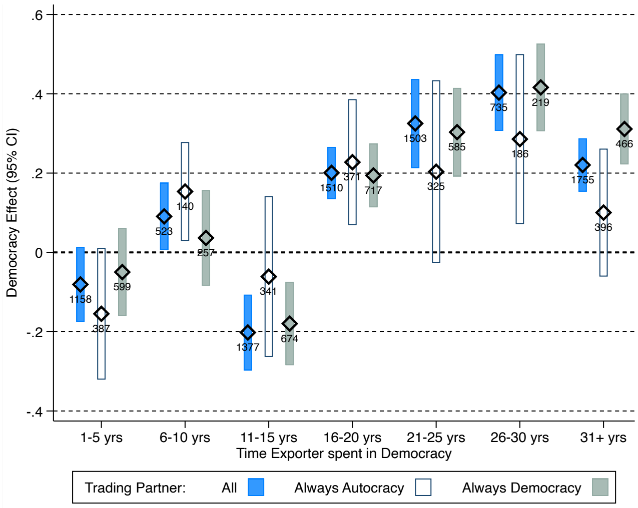
Figure C-2: Democratic Regime Change and Trade Flows — Alternative Democracy Dummies



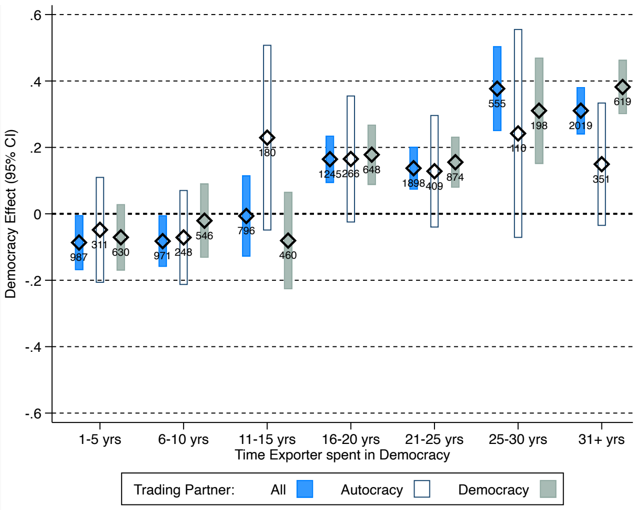
Notes: We use different definitions for our benchmark Liberal Democracy regime change — reprinted in panel (a): four versions, in panels (b) to (e) where we still use the Mean for the V-Dem LibDem index but subtract (add) 1/4 or 1/2 of a standard deviation to establish the democracy cutoff. In panel (e) we adopt the definition of Acemoglu et al (2019) which covers 1960-2010 (rather than 1950-2014). All plots have the same scale for the y-axis to ease comparison. For all other details see Figure C-1. Total estimates (full sample excluding zero-weight observations) amount to (a) 7,732 (b) 8,176, (c) 8,277, (d) 6,294, (e) 4,795, (f) 7,252.

Figure C-3: Democratic Regime Change and Trade Flows — Building Blocks of Liberal Democracy

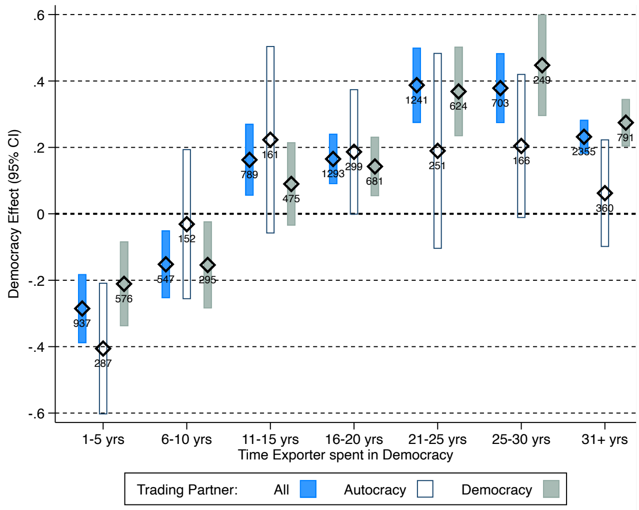
(a) Baseline: LibDem model (with Controls)



(b) Polyarchy model (with Controls)



(c) Liberal component Model (with Controls)



Notes: Total estimates (full sample excluding zero-weight observations) amount to (a) 7,732 (b) 9,364, (c) 8,702. Pseudo-Alpha tests reject the null of common expected factor loadings for the polyarchy (27.34, $p < 0.01$) and also for the liberal component (5.23, $p = 0.02$).

D Results: ATET Estimates

Table D-1: Average Treatment Effects and Pseudo-Alpha Test Statistics

Sample	(1) All	(2) Auto	(3) Demo	(4) Aut-Dem	(5) All	(6) Auto	(7) Demo	(8) Aut-Dem
Results in the main paper								
	(A) Baseline: (LibDem with Controls)				(B) Single Regime Change			
ATET	0.137***	0.093**	0.139***	0.175***	0.190***	0.169***	0.167***	0.291***
	[0.017]	[0.039]	[0.021]	[0.036]	[0.028]	[0.060]	[0.034]	[0.072]
Pairs	8568	2456	3804	2308	4806	1378	2305	1123
Pseudo-Alpha (p)	1.54 (.21)				1.78 (.18)			
	(C) Destination Regime Change				(D) Single Regime Change			
ATET	0.120***	0.162***	0.090***	0.231***	0.114***	0.265***	0.075**	0.163*
	[0.019]	[0.051]	[0.021]	[0.052]	[0.029]	[0.076]	[0.033]	[0.085]
Pairs	5948	1656	3248	1044	3308	880	1892	536
Pseudo-Alpha (p)	2.07 (.15)				0.36 (.55)			
	(E) Cold War Sample				(F) Post-Cold War Sample			
ATET	0.044*	0.093*	0.051**	-0.114	0.054***	0.080*	0.046*	0.074
	[0.025]	[0.048]	[0.026]	[0.074]	[0.020]	[0.044]	[0.024]	[0.054]
Pairs	2237	1157	766	314	5879	1608	3205	1066
Pseudo-Alpha (p)	1.52 (.22)				1.81 (.18)			
Results in the appendix								
	(G) Countries with Single Regime Change				(H) Countries with Multiple Regime Changes			
ATET	0.190***	0.169***	0.167***	0.291***	0.102***	-0.001	0.114***	0.150***
	[0.028]	[0.060]	[0.034]	[0.072]	[0.021]	[0.054]	[0.026]	[0.039]
Pairs	4806	1378	2305	1123	1969	520	706	743
Pseudo-Alpha (p)	1.78 (.18)				2.26 (.12)			
	(I) Baseline without Controls				(J) Baseline ignoring Zeros			
ATET	0.139***	0.090**	0.143***	0.180***	0.315***	0.267***	0.265***	0.458***
	[0.017]	[0.039]	[0.021]	[0.037]	[0.018]	[0.039]	[0.023]	[0.036]
Pairs	8596	2469	3810	2317	8791	2539	3888	2364
Pseudo-Alpha (p)	2.24 (.13)				34.79 (.00)			
	(K) Baseline: (LibDem with Controls)				(L) LibDem Cutoff: Mean – 1/2 SD			
ATET	0.137***	0.093**	0.139***	0.175***	0.186***	0.161***	0.187***	0.213***
	[0.017]	[0.039]	[0.021]	[0.036]	[0.018]	[0.052]	[0.022]	[0.035]
Pairs	8568	2456	3804	2308	9103	1542	4526	3035
Pseudo-Alpha (p)	1.54 (.21)				5.86 (.02)			
	(M) LibDem Cutoff: Mean – 1/4 SD				(N) LibDem Cutoff: Mean + 1/4 SD			
ATET	0.183***	0.148***	0.200***	0.192***	0.165***	0.143***	0.120***	0.358***
	[0.017]	[0.040]	[0.021]	[0.035]	[0.021]	[0.038]	[0.025]	[0.055]
Pairs	9087	2181	4182	2724	6893	2875	2594	1424
Pseudo-Alpha (p)	5.29 (.02)				1.18 (.28)			
	(O) LibDem Cutoff: Mean + 1/2 SD				(P) ANRR Democracy Dummy (1960-2010)			
ATET	0.238***	0.194***	0.198***	0.514***	0.274***	0.264***	0.254***	0.325***
	[0.023]	[0.038]	[0.028]	[0.067]	[0.020]	[0.048]	[0.026]	[0.041]
Pairs	5298	2601	1916	781	8092	1961	3881	2250
Pseudo-Alpha (p)	0.00 (.99)				0.83 (.36)			

Notes: The table presents robust mean estimates (average treatment effects on the treated computed using robust regression) for all specifications and subsamples considered in the paper: Figure 1 in the main text and Figures C-1 and C-2 in the Appendix. Columns indicate the destination (in (C) and (D) exporter) regime status, including the sub-sample of 'democratisers' (Aut-Dem). The Pseudo-Alpha test is for the null hypothesis that the coefficients on the cross-section averages sum to 1, implemented using a seemingly unrelated regression framework.

E The PPML-CCE-DID Estimator

In this Appendix Section, we take more time to motivate and develop our empirical implementation to capture monadic variables in a factor-augmented heterogeneous gravity model. The first two sub-sections are a recap of the modern structural gravity empirics, before section E.3 introduces heterogeneous gravity. Our combination of the PPML-CCE and recent heterogeneous difference-in-differences estimators is presented thereafter.

E.1 The Gravity Equation

We assume a gravity relationship in the panel in which bilateral trade of exporter i to destination market j at time t is given by[28]

$$X_{ijt} = \frac{Y_{it}}{\Omega_{it}} \frac{X_{jt}}{\Phi_{jt}} \phi_{jit} \qquad \text{where} \quad 0 \leq \phi_{ijt} \leq 1. \tag{3}$$

X_{ijt} is a trade flow from an exporter to a destination market, Y_{it} is the value of production for the exporter and X_{jt} the value of expenditure in the destination market j on all source countries — the latter two are typically proxied by GDP in the exporter and destination markets, respectively.[29] ϕ_{ijt} captures the 'bilateral accessibility' for destination j and exporter i: this contains trade costs between the two markets and any variable which may affect these, including time-variant and invariant, observed and unobserved factors.

A major development in gravity modelling over the past decade following the seminal contribution by Anderson and Van Wincoop (2003) is the recognition that the conditional trade between destination j and exporter i (the conditions being the 'bilateral accessibility') cannot be viewed in isolation from the set of opportunities open to importer j in sourcing goods from exporters other than i and the relative access exporter i has to destinations other than j.[30] The multilateral resistance variables for each actor in the exchange of goods are defined in terms of the bilateral accessibility-weighted exporter capabilities and importer characteristics respectively: exporter $i = 1, \ldots, N-1$ and for importer $j = 1, \ldots, N-1$, and $i \neq j$ let

$$\Omega_{it} = \sum_{\ell=-i} \frac{\phi_{i\ell t} X_{\ell t}}{\Phi_{\ell t}} \qquad \Phi_{jt} = \sum_{\ell=-j} \frac{\phi_{\ell jt} Y_{\ell t}}{\Omega_{\ell t}} \tag{4}$$

where $-j$ and $-i$ signify that these magnitudes are not defined in reflexive terms and thus exclude destination j and exporter i from the respective MRTs of the trading relationship between these two markets.

For our derivation of the empirical gravity model, we assume a stochastic version of equation (3)

$$X_{ijt} \quad = \quad \frac{Y_{it}}{\Omega_{it}} \frac{X_{jt}}{\Phi_{jt}} \phi_{ijt} \eta_{ijt} \tag{5}$$

where η_{ijt} is an error factor with $E[\eta_{ijt} | Y_{it}, \Omega_{it}, X_{jt}, \Phi_{jt}, \phi_{ijt}] = 1$.

A very general empirical equivalent to equation (5) allows for flexible unknown parameters on the observable mass and accessibility variables:

$$\begin{aligned} X_{ijt} \quad = \quad & exp[\beta^i_{ijt} \, ln(Y)_{it} + \beta^n_{ijt} \, ln(X)_{jt} + \gamma_{ijt} \, ln(\phi)_{ijt} \\ & + ln(\Omega)_{it} + ln(\Phi)_{jt}] \, \eta_{ijt}, \end{aligned} \tag{6}$$

where superscripts are used to identify the coefficient of exporter versus importer GDP/expenditure. This

[28] This exposition builds on the discussion in Santos Silva and Tenreyro (2006), and Yotov et al. (2016) for the gravity model, and Cameron and Trivedi (1998) for Poisson regression.

[29] If X_{ij} is merchandise trade then theory-consistency dictates Y_i to be gross production of traded goods (not simply value-added/GDP) and X_j the apparent consumption of goods, production plus imports minus exports (Head and Mayer, 2014).

[30] This network of dependencies is formalised by econometricians as the deviation from the assumption of 'cross-section weak dependence' (Andrews, 2005; Chudik, Pesaran, and Tosetti, 2011; Chudik and Pesaran, 2015b).

specification is the most general empirical model possible where all unknown parameters on observable variables (β^i, β^j, γ) vary at the pair level and over time. We demonstrate how this relates to the models employed in the existing literature by adopting parameter restrictions.

E.2 Model Restrictions and Pooled Estimation

A first restriction, adopted in the vast majority of studies, is to assume the gravity model estimates are fixed *across time*:

$$\begin{aligned} X_{ijt} &= exp[\beta^i_{ij}\, ln(Y)_{it} + \beta^n_{ij}\, ln(X)_{jt} + \gamma_{ij}\, ln(\phi)_{ijt} \\ &\quad + ln(\Omega)_{it} + ln(\Phi)_{jt}]\, \eta_{ijt}, \end{aligned} \tag{7}$$

A notable exception here is the study by Klasing, Milionis, and Zymek (2015) who allow for time-variation in three distinct regimes over their long panel from 1870 to 2005. In our analysis of the post-WWII period we first follow the bulk of the literature and estimate policy effects which are specified as time-invariant; later on we *partly* relax the assumption of fixed parameters over time by adopting 20-year rolling regression windows.

Conventionally, further restrictions in the panel gravity literature are to assume common parameters on the observable variables ($\beta^i_{ij} = \beta^i$, $\beta^n_{ij} = \beta^n$, $\gamma_{ij} = \gamma$). Pairwise fixed effects (δ_{ij}) are added to capture trade policy endogeneity (Baier and Bergstrand, 2007). Additionally including exporter-time (ω_{it}) and importer-time (ψ_{jt}) fixed effects can capture the MRTs (Hummels, 2001; Anderson and van Wincoop, 2003; Feenstra, 2004) — with implications for β^i and β^j (see below).

$$\begin{aligned} X_{ijt} &= exp[\beta^i\, ln(Y)_{it} + \beta^j\, ln(X)_{jt} + \gamma\, ln(\phi)_{ijt} + \delta_{ij} \\ &\quad + \omega_{it} + \psi_{jt}]\, \eta_{ijt}. \end{aligned} \tag{8}$$

Thus δ_{ij}, ϕ_{jt} and ω_{it} are the unknown parameters estimated on the various fixed effects in the reduced-form panel gravity model.

E.3 Heterogeneous Parameter Estimation

A practical difficulty arises if a more flexible specification for the observable variables such as that laid out in equation (7) on the one hand (β^i_{ij}, β^n_{ij}, γ_{ij}), and the aforementioned recommended practice to capture MRTs on the other are to be combined: for pairwise heterogeneity in the economic mass and policy variable parameters it is most convenient to estimate equation (7) *separately* for each country pair. However, we cannot include fixed effects for *all exporters and all importers* in an equation for a single importer-exporter pair, let alone interacted with time dummies. Existing studies in the literature which allow for trade policy heterogeneity employ interaction effects (e.g. Baier, Bergstrand and Clance, 2018) or maintain a set of fixed effects in a PPML model but use pair-specific trade policy dummies (Baier, Yotov and Zylkin, 2019). An alternative approach is to draw on recent insights from the panel time series literature (e.g. Pesaran, 2006; Bai, 2009) and to employ a multi-factor error structure to capture the unobserved MRTs.

First, we bring the error factor on the inside of the exponential function to capture all unobservables u_{ijt}:

$$\begin{aligned} u_{ijt} &= \delta_{ij} + \omega_{it} + \psi_{jt} + ln(\eta)_{ijt} \\ &\equiv \delta_{ij} + \omega_{it} + \psi_{jt} + \varepsilon_{ijt}. \end{aligned} \tag{9}$$

Next, the dimensionality problem of dealing with a large number of unknown parameters (the number of ω_{it} and ϕ_{jt} swiftly add up to thousands of directional-time dummies) in a country-pair equation with about 70 post-WWII time series observations can be solved by imposing more structure on these unobservables. In the macro panel econometric literature it is widely acknowledged that a small number of common factors with heterogeneous factor loadings can represent large datasets of macro variables. For instance, in the forecasting literature Stock and Watson (2002) have shown that 149 macroeconomic time series can be reduced to two or three principal components. Furthermore, Bai (2009) discusses several macro-, microeconometric and finance applications where common factors can be employed to model unobserved time-varying heterogeneity in a tractable way.[31]

In the case at hand, we posit that the economic mass and accessibility variables along with the trade flows and MRTs are all driven by a small number of common factors with heterogeneous factor loadings across country pairs. Thus, we argue that a small number of unobserved common factors $\boldsymbol{f}_t$, each with country pair-specific factor loadings $\boldsymbol{\varphi}_{ij}$, can account for the evolution of trade flows, GDP, etc. In the notation introduced above:

$$\begin{aligned} u_{ijt} &= \delta_{ij} + \omega_{it} + \psi_{jt} + \varepsilon_{ijt} && (10) \\ &\approx \delta_{ij} + \boldsymbol{\varphi}_{ij}' \boldsymbol{f}_t + \varepsilon_{ijt}, && (11) \end{aligned}$$

where an 'approximate factor structure' is represented by a set of common factors $\boldsymbol{f}$, and the associated factor loadings $\boldsymbol{\varphi}$.

The insight gained in the recent panel time series literature from this setup in the *linear* regression case is that the unobserved common factors can be captured by observables, either via principal component analysis (Bai, 2009) or using cross-section averages of the dependent and independent variables (Pesaran, 2006). We briefly develop the latter approach (the 'common correlated effects' or CCE estimator) and provide a mathematical indication of the intuition at play — this is for ease of illustration since it will only be a small step in terms of implementation from the linear model to a generalised linear one (Boneva and Linton, 2017).

For simplicity we assume a double-index of t for the time series and i for the cross-section — we can think of the latter as a placeholder for the country pair as the unit of analysis like in the gravity model, i.e. $i \equiv ij$. Let

$$\begin{aligned} y_{it} &= \alpha_i + \beta_i x_{it} + u_{it} \qquad u_{it} = \lambda_i f_t + \varepsilon_{it} && (12) \\ x_{it} &= \delta_i + \phi_i f_t + \varrho_i g_t + e_{it}, && (13) \end{aligned}$$

where ε and e are white noise processes. This setup indicates that the regressor x is driven by the same common factor f as the dependent variable y, albeit with different parameters.[32] In addition there are some factors $\boldsymbol{g}$ which only drive x but not y. This setup is standard in the macro panel literature and we refer to the studies in footnote 31 for details on factor evolution, parameter distributions, etc. It is clear from equations (12) and (13) that x is endogenous and that failing to account for the presence of the unobserved common factors will lead to omitted variable bias.[33]

Pesaran's (2006) approach posits that the unobserved common factor f can be captured by the cross-

[31] The multifactor error structure has been applied to capture country-specific time-varying total factor productivity (Eberhardt and Presbitero, 2015), time-varying absorptive capacity (De Visscher et al, 2020), or knowledge spillovers in the analysis of sector-level production functions augmented with sectoral R&D stock (Eberhardt, Helmers, and Strauss, 2013).

[32] In a setup with multiple factors we can instead assume that only a subset of $\boldsymbol{f}$ 'overlaps' between the two equations.

[33] Solving the x equation for f and plugging this into the y equation yields

$$\begin{aligned} y_{it} &= \alpha_i + \beta_i x_{it} + \lambda_i \phi_i^{-1}(x_{it} - \delta_i - \psi_i g_t - e_{it}) + \varepsilon_{it} && (14) \\ &= \alpha_i - \lambda_i \phi^{-1} \delta_i + (\beta_i + \lambda_i \phi_i^{-1}) x_{it} - \lambda_i \phi_i^{-1} \psi_i g_t - \lambda_i \phi_i^{-1} e_{it} + \varepsilon_{it} \\ &= \eta_i + \theta_i x_{it} + \nu_{it}, && (15) \end{aligned}$$

where in the final line we reparameterise. Crucially, unless $\lambda_i \phi_i^{-1} = 0$ we can see that β_i is unidentified. The asymptotic bias will be a function of the (relative) 'strength' of the factors in their impact on y and x in panel member i.

section averages of y and x provided the cross-section dimension of the panel is not too small.[34] A simple algebraic derivation can provide intuition for the mechanism at work: take the cross-section average of equation (12) and solve it for the common factor f:

$$\bar{y}_t \quad = \quad \bar{\alpha} + \bar{\beta}\bar{x}_t + \bar{\lambda}f_t \quad \Leftrightarrow \quad f_t = \bar{\lambda}^{-1}(\bar{y}_t - \bar{\alpha} - \bar{\beta}\bar{x}_t), \tag{16}$$

where the bars indicate cross-section averages and the error term disappears since $\bar{\varepsilon} = 0$ by assumption. Next, plug the expression for f back into the original equation (12)

$$\begin{aligned} y_{it} &= \alpha_i + \beta_i x_{it} + \lambda_i \bar{\lambda}^{-1}(\bar{y}_t - \bar{\alpha} - \bar{\beta}\bar{x}_t) + \varepsilon_{it} & (17)\\ &= \alpha_i - \lambda_i \bar{\lambda}^{-1}\bar{\alpha} + \beta_i x_{it} + \lambda_i \bar{\lambda}^{-1}\bar{y}_t - \lambda_i \bar{\lambda}^{-1}\bar{\beta}\bar{x}_t + \varepsilon_{it}, \\ y_{it} &= \varpi_i + \beta_i x_{it} + \zeta_i \bar{y}_t + \vartheta_i \bar{x}_t + \varepsilon_{it}, & (18)\end{aligned}$$

where we reparameterize in the last line. Thus the unobserved common factor f can be captured by the cross-section averages of y and x, while the heterogeneous impact of f across i can be captured by estimating equation (18) separately for each panel member — the principle extends to multiple factors. A Mean Group estimator following Pesaran and Smith (1995) captures the central tendency of the panel and provides a convenient comparison with alternative pooled empirical models:

$$\hat{\beta}^{MG} = N^{-1}\sum_{i=1}^{N}\hat{\beta}_i \tag{19}$$

Inference for the Mean Group estimates is based on a simple nonparametric variance estimator (Pesaran and Smith, 1995; Pesaran, 2006):[35]

$$\widehat{\Omega}^{MG} = \frac{1}{N-1}\sum_{i=1}^{N}(\hat{\beta}_i - \hat{\beta}^{MG})(\hat{\beta}_i - \hat{\beta}^{MG})'. \tag{20}$$

Boneva and Linton (2017) extend the above setup from the linear model to a setting where the outcome variable is discrete. They show that the CCE approach can be applied to a probit model by including only the cross-section averages of the observed regressors: under the assumption that the unobserved factors are contained in the span of the cross-section averages of the regressors they derive asymptotic results for the large T, large N case as well as the consistency and asymptotic normality of the Mean Group estimator of the individual-specific estimates.

The same principle can be applied to a generalised linear model: in our case, we begin by assuming the exponential mean function incorporating a multi-factor error structure[36]

$$E[y_{it}|x_{it}, \boldsymbol{f}_t] = \mu_{it} = exp[\alpha_i + \beta_i x_{it} + \boldsymbol{\lambda}_i' \boldsymbol{f}_t]. \tag{21}$$

The Poisson Maximum Likelihood (PML) estimator assumes that the distribution of y given x (and the factors) is Poisson (i.e. a count variable), but it is widely recognised that the data generating process is not required to be Poisson for this estimator to be consistent (Cameron and Trivedi, 1998; Santos Silva and Tenreyro, 2006), in which case it is referred to as a Poisson Pseudo Maximum Likelihood (PPML) estimator. Our implementation uses the estimator in the single time series, namely of the pair-wise gravity equation with exponential mean

[34] Additional robustness of this approach to nonstationary factors, structural breaks, and additional spatial dependence, among other aspects, is discussed in Kapetanios, Pesaran, and Yamagata (2011), Chudik, Pesaran and Tosetti (2011), Chudik and Pesaran (2015a) and Westerlund (2019).

[35] In practice we follow the standard in the literature and employ robust means (Hamilton, 1992) to reduce the effect of outliers.

[36] We thank Lena Boneva and Oliver Linton for sharing the rough derivations for a CCE-augmentation in this more general case.

function, where the common factors are replaced by cross-section averages of the regressors, which we refer to as PPML-CCE:

$$\begin{aligned}\mu_{ijt} &= exp[\alpha_{ij} + \beta^i_{ij}\, ln(Y)_{it} + \beta^n_{ij}\, ln(X)_{jt} + \gamma_{ij}\, ln(\phi)_{ijt} \\ &\quad +\delta_{ij}\, \overline{ln(Y)}_t + \kappa_{ij}\, \overline{ln(\phi)}_t] + \epsilon_{ijt} \qquad \forall\, ij,\, j \neq i.\end{aligned} \tag{22}$$

In our case, the accessibility term (any variable that may affect bilateral trade costs) is replaced by dummies for democratic regime change in i and j respectively along with dyadic dummies for free trade agreements and currency unions (in robustness checks we omit the latter two since they may be viewed as 'bad controls').[37] In practice, where exporter and destination mass are proxied using their respective GDP, the cross-section averages are identical, $\overline{ln(Y)}_t = \overline{ln(X)}_t$. In our difference-in-differences implementation below these come from distinct control samples and are thus separately identified.

Our empirical estimates below are thus based on cross-section average-augmented time series PPML regressions at the country-pair level which are subsequently averaged following the Mean Group principle; inferential statistics for the PPML-CCE Mean Group estimate are computed using the variance estimator in (20): Boneva and Linton (2017) have shown the simple nonparametric variance estimator still applies in the generalised linear model setup.

Estimation of the PPML-CCE model can be seen as an improvement on current practice for two reasons. First, the presence of common factors, with heterogeneous loadings across country pairs, offers a flexible way to account not only for the MRTs, but also so-called globalisation effects, and other forms of spatial dependence. Second, using common factors as proxies for spatial dependence allows the estimation of country-specific (monadic) variables — thus in addition to dyadic determinants such as those contained in $ln(\phi)_{ijt}$, we can also introduce country-specific variables of interest.

E.4 A Difference-in-Differences Gravity Model

Our exposition so far has focused on heterogeneous gravity without giving too much thought to the empirical practices in the democracy-growth literature, where economists favour using *binary* indicators for democracy versus autocracy (Papaioannou and Siourounis, 2008; Acemoglu et al, 2019; Eberhardt, 2022; Boese-Schlosser and Eberhardt, 2023, 2024). We now introduce a combination of the PPML-CCE model with a factor-augmented difference-in-differences estimator following Chan and Kwok (2022): let D_{it} be a dummy for democratic regime change equal to 0 when country i is in autocracy and 1 if they are in democracy — we refer to the shift from 0 to 1 as 'treatment'. The Chan and Kwok (2022) principal component DID (PCDID) estimator estimates the causal effect of a treatment in country-specific regressions of treated countries only. While this appears to capture only the first of the difference-in-differences, the comparison with the control sample (in our case those countries which never experienced democratic regime change) is achieved by the inclusion of unobserved common factor proxies extracted from the control sample of never-democratisers. Crucially, this heterogeneous difference-in-differences estimator does not require the 'parallel trends' assumption to hold, but relies on a much weaker assumption of 'weak parallel trends' between treated and control samples — we provide more details in the next section where we develop a suitable test in the following subsection.

In the context of democratic regime change and the PPML-CCE gravity model of trade flows, we make some adjustments to the way the Chan and Kwok (2022) PCDID is implemented: first, we do not *estimate* the unobserved common factors from auxiliary regressions in the control sample of never-democratisers. Extracting several principal components via PCA from the residuals of a PPML regression is conceptually difficult (given

[37] We know that the dependent variable, conditional on the regressors, can be heteroskedastic, and serially correlated and that the PPML estimator maintains its consistency in the presence of serial correlation if the exponential conditional mean is correctly specified (Cameron and Trivedi, 1998: 226) — the latter is the requirement for any PML or PPML estimator.

that these are not linear models) and in practice raises concerns over how many estimated factors to include. Instead, we use cross-section averages[38] of GDP and population in line with Boneva and Linton (2017). Second, these cross-section averages are constructed from the respective control samples of (i) exporters and (ii) destinations: since the trade flow panel is not symmetric (i may be exporting to j but j does not export to i) these two cross-section averages are not identical.

Our PPML-CCE-DID estimator is then implemented as follows:

$$\begin{aligned} \mu_{ijt} \quad = \quad & exp[\alpha_{ij} + \gamma^1_{ij}\, \text{FTA}_{ijt} + \gamma^2_{ij}\, \text{Common Currency}_{ijt} + \theta_{ij} D_{it} + \eta_{ij} D_{jt} \\ & + \delta^i_{ij}\, \overline{ln(Y)}^i_t + \kappa^i_{ij}\, \overline{ln(Pop)}^i_t + \delta^j_{ij}\, \overline{ln(Y)}^j_t + \kappa^j_{ij}\, \overline{ln(Pop)}^j_t + \epsilon_{ijt}] \qquad \forall\, ij,\, j \neq i. \end{aligned} \tag{23}$$

We have pair fixed effect (α_{ij}) and the familiar dyadic variables for FTA and common currency union. The two democracy dummies (D_{it} for exporter i and D_{jt} for destination j) are accompanied by two sets of cross-section averages, which are constructed from the exporter i and destination j control samples of countries which never democratised, respectively. We do not include cross-section averages of the indicator variables in the first line of equation (23) since this practice is likely to induce severe bias in the estimation equation (Juodis et al, 2021). Our coefficient of interest is θ_{ij}, which is estimated (separately) in each trade flow equation of country i with each of its trade partners j — as is standard in the heterogeneous DID literature, we estimate equation (23) only for equations where either country i or country j experienced regime change.

We cannot estimate the democracy effect in country-pairs where trade flows remained zero throughout the sample period, whether they experienced regime change or not: this is necessitated by our heterogenous regression implementation. We do however include the GDP and population variables for these country pairs in the computation of the cross-section averages. More worryingly, excluding country pairs where the exporter or importer did experience regime change but nevertheless, trade between the pair never materialised, will likely induce an upward (selection) bias in our results. We therefore adjust the model to include cross-section averages from those country pairs where the exporter or importer country did experience regime change but where trade flows remained zero. Our final empirical implementation is thus:

$$\begin{aligned} \mu_{ijt} \quad = \quad & exp[\alpha_{ij} + \gamma^1_{ij}\, \text{FTA}_{ijt} + \gamma^2_{ij}\, \text{Common Currency}_{ijt} + \theta_{ij} D_{it} + \eta_{ij} D_{jt} \\ & + \delta^i_{ij}\, \overline{ln(Y)}^i_t + \kappa^i_{ij}\, \overline{ln(Pop)}^i_t + \delta^j_{ij}\, \overline{ln(Y)}^j_t + \kappa^j_{ij}\, \overline{ln(Pop)}^j_t \\ & + \delta^{i0}_{ij}\, \overline{ln(Y)}^{i0}_t + \kappa^{i0}_{ij}\, \overline{ln(Pop)}^{i0}_t + \delta^{j0}_{ij}\, \overline{ln(Y)}^{j0}_t + \kappa^{j0}_{ij}\, \overline{ln(Pop)}^{j0}_t + \epsilon_{ijt}] \qquad \forall\, ij,\, j \neq i, \end{aligned} \tag{24}$$

where the superscript $i0$ ($j0$) indicates the cross-section averages or associated parameters from exporter (destination) countries which did experience regime change but where trade flows between i and j remained zero throughout the sample period. This specification is even more demanding on the data, with 13 parameters to be estimated in each country-pair regression.

[38] Eberhardt (2022) adopts the same strategy in a standard (monadic) panel dataset to study democracy and growth.