European Journal of Political Economy 37 (2015) 49–63

Contents lists available at ScienceDirect

European Journal of Political Economy

journal homepage: www.elsevier.com/locate/ejpe

Trade openness and bigger governments: The role of country size revisited

Michael Jetter [a,*], Christopher F. Parmeter [b]

[a] Department of Economics, Universidad EAFIT, IZA, Carrera 49 #7 Sur-50, Avenida Las Vegas, Medellín, Colombia
[b] Department of Economics, University of Miami, United States

ARTICLE INFO

Article history:
Received 16 May 2014
Received in revised form 31 October 2014
Accepted 1 November 2014
Available online 12 November 2014

JEL classification:
C21
C23
F62
H10
H50

Keywords:
Government size
Trade openness
Country size
Panel data
Penn world table

ABSTRACT

This paper revisits the question of why more open countries tend to have bigger governments. We replicate successfully the main results of Ram (2009), who rejects the role of country size as an omitted variable. However, several extensions advise against a hasty conclusion: The results differ substantially depending on the data source used, the timeframe considered, the countries selected, and the way variables are measured. Specifically, we employ newer versions of the Penn World Table (PWT 7.1 and 8.0), allowing us to both extend the number of observations and the timeframe. We find evidence for the claim that smaller countries do indeed have bigger governments, especially when using the PWT 8.0 data, and Ram (2009) findings might be driven by the specific dataset used (PWT 6.1) and the countries included in that sample. Finally, we also conduct quantile regression analyses to pin down at which point of the distribution the suggested relationships come out.

1. Introduction

Over the past 50 years, the world has experienced a prodigious rise in international trade relationships. Trade openness of the average country, measured as exports plus imports as a share of GDP, has increased from 54.4% in 1960 to 87.3% in 2010.[1] As globalization has progressed, so has the need to understand the consequences from trade. In the public economics literature, Cameron (1978) is among the first to note that trade openness in OECD countries is positively related to subsequent government size. Since the size of governments may carry macroeconomic consequences, for example impeding economic growth (see Barro and Lee, 1994; Barro, 2001; Afonso and Furceri, 2010; Bergh and Henrekson, 2011), it is important to understand which factors shape the extent of the public sector.

Twenty years later, Rodrik (1998) suggests a potential causality between trade openness and government size. In his explanation, trade openness is associated with increased uncertainty, which in turn leads citizens to request stronger public safety nets. Thus, a higher degree of openness should lead to bigger governments in per capita terms. Following Rodrik (1998), numerous papers challenge this "compensation hypothesis" (e.g. Liberati, 2007; Benarroch and Pandey, 2008; Dreher et al., 2008). However, there exist

* Corresponding author. Tel.: +57 317 433 7003.
E-mail addresses: mjetter7@gmail.com (M. Jetter), cparmeter@bus.miami.edu (C.F. Parmeter).
URL: http://www.michaeljetter.com (M. Jetter).

[1] Numbers derived from the Penn World Table version 7.1, comparing 112 countries, which show data for trade openness in 1960 and 2010.

http://dx.doi.org/10.1016/j.ejpoleco.2014.11.001

few alternative theories on the observed positive link between trade openness and the size of government. One of the strongest explanations is developed by Alesina and Wacziarg (1998) (AW hereafter), who argue that the size of a country influences both variables: Smaller countries tend to have bigger governments in per capita terms (less economies of scale), but are also more open to international trade, naturally owed to physical limitations in their domestic economies. Both Rodrik (1998) and AW find empirical evidence lending support to their theories.

A decade later, Ram (2009) revisits both competing theories in a panel data framework using a more expansive and updated dataset than found in either Rodrik (1998) or AW. Investigating these two competing theories on government size with panel data hinges critically in the presence of unobserved country specific heterogeneity, which is likely to be present. Using Penn World Table data version 6.1 (PWT 6.1) from 1960 to 2000, Ram's (2009) results support the openness to trade theory, whereas Alesina and Wacziarg's (1998) theory is not confirmed by the data when controlling for unobserved country and time specific heterogeneity.

This paper replicates the main results of Ram (2009) and extends his empirical analysis in several dimensions. We consider data source and timeframe extensions, as well as the set of countries selected, the form of measurement for the key variables, and alternative features of the conditional distribution aside from the conditional mean (quantiles). The use of alternative data sources represents an attractive avenue to assess the robustness of Ram's (2009) insights given that different updates of the PWT have been criticized (see Breton, 2012) and provide more data in both the country and time dimensions. Checking the robustness of empirical results across various updates of the PWT within the cross-country growth literature has unveiled serious concerns regarding many key insights from older versions (Johnson et al., 2013), drawing into question the robustness of standard results. To date, these analyses center around the GDP per capita estimates and the resulting growth literature, for example by Atherton et al. (2011) and Ciccone and Jarocinski (2010).[2]

The results stemming from our extensions of Ram (2009) provide another avenue beyond cross-country growth for which the robustness of key empirical results is not immune to alternative updates of the Penn World Table. Finally, quantile analyses are able to trace out a better understanding as to where exactly the suggested relationships are present. Taken together, our extensions present evidence that the main findings of Ram (2009) are not as clear cut as one might hope. Our findings demonstrate that changes in the timeframe, dataset, and sample countries, along with focusing on conditional quantiles (as opposed to the conditional mean), demonstratively provide less clear support for Rodrik (1998) and Ram (2009). This implies that there are still theoretical links which can be exploited to connect the works of AW and Rodrik (1998). In general, the hypothesis put forth by AW, namely that country size is driving the observed positive link between trade openness and government size, cannot be dismissed.

The following section outlines the methodology and data of our analysis. Section 3 presents our findings and Section 4 concludes with a brief summary of our results.

2. Empirical strategy

2.1. Methodology

In order to estimate the effects of country size and trade openness on government size, Ram (2009) estimates three main equations of interest for country *i* at year *t*:

$$LGOV_{it} = a_0 + a_1 LSIZE_{it} + a_2 LRY_{it} + \alpha_{1i} + \lambda_{1t} + u_{1it}, \tag{1}$$

$$LOPEN_{it} = b_0 + b_1 LSIZE_{it} + b_2 LRY_{it} + \alpha_{2i} + \lambda_{2t} + u_{2it}, \tag{2}$$

and

$$LGOV_{it} = d_0 + d_1 LOPEN_{it} + d_2 LRY_{it} + \alpha_{3i} + \lambda_{3t} + u_{3it}, \tag{3}$$

where *LGOV* and *LSIZE* measure the size of government and the size of the country. *LOPEN* represents trade openness of the country and *LRY* measures the wealth of the country. The natural logarithm is applied to all variables for comparability of results, and is denoted by an *L* in front of each variable. Country and time specific heterogeneity are captured by α_{ji} and λ_{jt}, while u_{jit} represents the idiosyncratic error term for $j = 1, 2, 3$.

Eqs. (1) and (2) represent the empirical counterpart to the theoretical insights of AW, whereas Eq. (3) is the empirical model of Rodrik (1998). AW pointed out that one may witness a positive coefficient in Eq. (3), but this may not necessarily reflect the causal impact of trade openness on government size. Specifically, smaller countries both tend to select bigger governments in relative terms and be more open to international trade relationships. The first observation relies on a basic economy of scale argument, since a variety of public institutions and services (e.g., national security) require substantial setup costs, but the cost per capita only increases marginally with a bigger population. The second relationship stems from the observation that smaller countries naturally exhibit smaller domestic markets and international trade relationships alleviate this shortcoming. Hence, AW's theory was one of an omitted variable – namely country size – that, when accounted for, explained the finding in Rodrik (1998).

[2] The development of PWT 8.0 focuses on addressing these issues with a strong emphasis on the calculation of PPPs when deriving GDP per capita. See Feenstra et al. (2009) and the PWT 8.0 user guide (Feenstra et al., 2011).

Table 1
Number of countries and observations (in parentheses under the number of countries).

	Annual	5-year mean	10-year mean
1960–2000			
PWT 6.1	154	150	135
	(5117)	(1027)	(509)
PWT 7.1	189	189	176
	(6373)	(1303)	(670)
PWT 8.0	167	167	167
	(5780)	(1181)	(619)
1960–2010			
PWT 7.1	190	190	190
	(8274)	(1684)	(874)
PWT 8.0	167	167	167
	(7447)	(1515)	(786)
Balanced 1960–2000			
PWT 6.1	98	112	115
	(4018)	(896)	(460)
PWT 7.1	111	113	159
	(4551)	(904)	(636)
PWT 8.0	107	109	143
	(4387)	(872)	(571)
Balanced 1960–2010			
PWT 7.1	111	113	159
	(5661)	(1130)	(795)
PWT 8.0	107	109	143
	(5457)	(1090)	(715)

Country wealth, as proxied by income per capita (LRY_{it}), is included for consistency with the empirical works of Rodrik (1998), AW, and Ram (2009). The underlying logic of controlling for income levels is based on Wagner's Law, which predicts that richer societies construct bigger governments in relative terms. For recent findings on the persistence of Wagner's Law, one might consider Akitoby et al. (2006), who consider developing countries, or Durevall and Henrekson (2011) for a long term study on Sweden and the United Kingdom. Ram (2009) estimates all three equations as pooled panels for consistency with the empirical work performed in each paper, but then focuses on the differences that arise when accounting for country and time specific heterogeneity within a fixed effect framework.[3]

In addition to using annual data, Ram (2009) re-estimates the models in Eqs. (1) to (3) using 5 and 10 year averages of each of the variables. These extensions serve as useful robustness checks for the main results, since averaging over time addresses potential problems arising from measurement errors and alleviates the effects from business cycles.

2.2. Data

To estimate the models in Eqs. (1) to (3), Ram (2009) uses an unbalanced panel of 154 countries over the 1960–2000 period taken from Penn World Table 6.1 (PWT 6.1, Heston et al., 2002). Ram (2009) proxies government size with the share of government consumption in GDP, country size with the total population of a country, and openness with the ratio of exports plus imports to GDP. In PWT 6.1, these variables are *cg*, *pop* and *openc*, respectively. Further, given the diverse range of countries in the Penn World Table, income per capita is measured in international dollars at constant prices, *rgdpch*. After replicating the results from Ram (2009), we extend the analysis in several dimensions, as explained in detail below. Table 1 provides the number of countries and observations across the 10 different datasets deployed in our empirical analysis.

In general, we choose the same variables from all three data sources. However, an interesting difference can be found in the variable definitions across our datasets. In this context, Breton (2012) and Johnson et al. (2013) point out that seemingly the same variables can differ remarkably across versions of the Penn World Tables. For instance, the PWT measurement for government size is labeled *Government Consumption Share of PPP Converted GDP Per Capita at current prices* [cgdp], (%), and Breton (2012) notices that a change in benchmark prices leads to substantial differences of the same variable from the PWT 6 series to the PWT 7 series. In fact, the description of PWT versions 7.0 and 7.1 acknowledges the changes from previous versions for the main expenditure groupings of government current expenditure (variable *cg*).[4]

Set aside the differences in benchmark prices used to calculate variables, the PWT 7.1 defines all variables in the same way as PWT 6.1 (*cg*, *pop*, *openc*, and *rgdpch*). In the case of PWT 8.0, government consumption and population size are measured as csh_g and *pop*, whereas we estimate trade openness by $(v_x + v_m)/v_g dp$, as recommended by the PWT 8.0 user guide. Similarly, *rgdpch* is calculated as *rgdpe*/*pop*. However, the major difference to previous PWT versions consists in measuring openness in real terms, rather than nominal

[3] Ram (2009) also considers using random effects, but Hausman tests unequivocally reject this framework.
[4] Also see page 27 of the *Description of PWT 7.0 and 7.1* link on the PWT website (https://pwt.sas.upenn.edu/php_site/pwt_index.php).

terms (Alcalá and Ciccone, 2004). In summary, any differences derived that are owed to the dataset employed, and not to timeframe or sample countries included, could potentially be associated with changes in measurement techniques, such as the usage of different benchmark prices or different variable definitions. For a deeper discussion on the differences between the PWT 8.0 and its predecessors, we recommend the PWT 8.0 user guide, page 32ff.

2.3. The replication and methodological issues

We start by replicating the main findings of Ram (2009), specifically his Tables 1–3. In each of these tables he estimates one of the Eqs. (1) to (3), treating the data first as a pooled cross section and then exploiting the full panel structure. Further, in each setting he uses annual, as well as 5 and 10 year averaged data. This produces a total of 18 sets of estimates (6 for each Eq. 3 using pooled OLS, and 3 using fixed effect estimation).

We are able to successfully replicate the results of Tables 1–3 of Ram (2009). However, several challenges are worth mentioning prior to presenting the replicated estimates and robust t-statistics. First, given the panel nature of the data and the different time averaging, several countries are dropped from the analysis. Ram (2009) never explicitly details these countries. We note that PWT 6.1 contains a total of 168 countries. Initially, 14 countries either do not possess useable data or only have a single year of data, making them poor candidates for inclusion in a panel data framework. These countries are Bahrain, the Bahamas, Bermuda, Bhutan, Djibouti, Eritrea, Kuwait, Laos, Mongolia, Oman, Qatar, Saudi Arabia, Sudan, and Turkmenistan.

Switching to the 5 year averaging, another four countries are excluded for having only a single, averaged observation. We note here that Ram (2009) reports in his notes to Tables 1–3 that progressing from the annual to 5 year averaging setup results in the loss of three countries. However, we find four countries to be omitted in addition to the previously mentioned 14: Georgia, Puerto Rico, Swaziland, and Tajikistan. Finally, another 15 countries are dropped for similar reasons when migrating from 5 year averaging to 10 year averaging. These 15 countries are Albania, Armenia, Azerbaijan, Bulgaria, Cambodia, Croatia, Estonia, Kazakhstan, Kyrgyzstan, Lithuania, Macedonia, Malta, Moldova, Russia, and Uzbekistan.

Another complication with replicating the main results in (Ram, 2009) is the fact that his 5 and 10 year averages are not traditional 5 and 10 year averages. Instead, his averaging uses the quintennial and decadial years in overlapping average observations. For instance, in the 5 year averages the year 1970 is used in the 1965 average $\left(\overline{z}_{i65} = (1/6)\sum_{t=1965}^{1970} z_{it}\right)$ as well as in the 1970 average $\left(\overline{z}_{i70} = (1/6)\sum_{t=1970}^{1975} z_{it}\right)$. Thus, his 5 and 10 year averages are actually 6 and 11 year averages. All regression results displayed throughout the paper follow this methodology of using 6 and 11 year averages for consistency purposes with Ram (2009). However, all derived results are qualitatively identical when using proper 5 and 10 year averages. We now move to describing our several extensions.

2.4. Extensions

First, we consider different data sources in re-estimating Eqs. (1) to (3) using PWT versions 7.1 and 8.0. To remain as close to Ram (2009) as possible, we start by keeping the same timeframe, namely 1960–2000, but include all available countries. Even when using the same 1960–2000 timeframe as Ram (2009), Table 1 shows that our two alternative datasets incorporate substantially more countries and annual observations. In the case of PWT 7.1, we see almost a 25% increase in the annual sample size. These differences in sample size are even more pronounced when considering the 5 and 10 year averaged datasets. The PWT 8.0 dataset does not lose a single country to 5 or 10 year averaging. Note also that PWT 7.1, published in 2012, actually produces more observations than PWT 8.0, published one year later.

Second, we take advantage of longer timeframes covered in the more recent datasets (PWT 7.1 and PWT 8.0). As both data sources offer information for 10 more years, we re-estimate Eqs. (1) through (3) using data covering 1960–2010. One advantage from using additional data is that we expect our estimates of the underlying relationships to be more precise given the inclusion of more data. However, we recuse ourselves from advocating on behalf of a "best" dataset. While it might be tempting to advocate on behalf of PWT 8.0, this is a naïve postulate if the argument is based on the vintage. When making international comparisons, no "best" dataset exists (Neary, 2004), and, as noted by Johnson et al. (2013), newer is not necessarily better. The root cause of problems associated with international comparisons is that different baskets of goods are produced at different points in time by different countries and no singular approach to making these baskets comparable exists. Given our inability to have an unambiguously preferred dataset, we simply mention that to adequately adjudicate between the competing insights of Rodrik (1998) and AW, robustness to data revisions should be a minimum requirement.

Table 2
Number of countries and observations that overlap with PWT 6.1 (1960–2000).

	Annual	5-year mean	10-year mean
PWT 6.1	154	150	135
	(5117)	(1027)	(509)
PWT 6.1–PWT 7.1	154	150	132
	(5085)	(1022)	(503)
PWT 6.1–PWT 8.0	143	140	125
	(4743)	(952)	(471)

Table 3
Replication and extensions of Ram (2009, Table 1). Dependent variable is the logarithm of government spending, while all coefficient estimates displayed are for the logarithm of population. All regressions control for the logarithm of GDP per capita.

	OLS			Two-way fixed effects		
	Annual	5-year mean	10-year mean	Annual	5-year mean	10-year mean
1960–2000						
PWT 6.1	−0.08***	−0.08***	−0.09***	0.48***	0.44***	0.42***
	(−17.43)	(−8.62)	(−6.67)	(10.84)	(4.95)	(3.67)
PWT 7.1	−0.12***	−0.12***	−0.12***	−0.00**	0.02	0.04
	(−30.71)	(−14.44)	(−10.66)	(−0.07)	(0.34)	(0.47)
PWT 8.0	−0.05***	−0.05***	−0.05***	−0.17***	−0.12	−0.10
	(−14.35)	(−6.32)	(−4.30)	(−4.75)	(−1.63)	(−1.11)
1960–2010						
PWT 7.1	−0.12***	−0.12***	−0.12***	−0.07***	−0.07	−0.07
	(−35.28)	(−16.53)	(−12.17)	(−2.69)	(−1.22)	(−0.97)
PWT 8.0	−0.06***	−0.06***	−0.05***	−0.37***	−0.30***	−0.27***
	(−18.84)	(−8.13)	(−5.77)	(−11.64)	(−4.41)	(−3.14)
Balanced 1960–2000						
PWT 6.1	−0.03***	−0.04***	−0.04***	0.55***	0.53***	0.55***
	(−5.63)	(−4.15)	(−3.10)	(11.51)	(5.51)	(4.43)
PWT 7.1	−0.10***	−0.10***	−0.12***	0.01	0.01	0.08
	(−16.75)	(−7.36)	(−10.58)	(0.38)	(0.17)	(1.05)
PWT 8.0	−0.04***	−0.03***	−0.05***	−0.30***	−0.25***	−0.05
	(−8.16)	(−3.58)	(−4.62)	(−7.50)	(−2.92)	(−0.53)
Balanced 1960–2010						
PWT 7.1	−0.09***	−0.08***	−0.12***	−0.05	−0.07	−0.01
	(−17.24)	(−7.65)	(−11.87)	(−1.61)	(−1.11)	(−0.09)
PWT 8.0	−0.04***	−0.04***	−0.06***	−0.56***	−0.54***	−0.36***
	(−9.68)	(−4.27)	(−5.99)	(−14.26)	(−6.23)	(−3.88)
Using consistent countries, 1960–2000						
PWT 6.1–PWT 7.1	−0.08***	−0.09***	−0.09***	0.51***	0.51***	0.53***
	(−17.69)	(−8.72)	(−6.80)	(22.17)	(10.42)	(8.82)
PWT 7.1–PWT 6.1	−0.11***	−0.11***	−0.11***	0.16***	0.16***	0.17***
	(−22.52)	(−10.15)	(−7.69)	(7.95)	(4.13)	(3.26)
PWT 6.1–PWT 8.0	−0.08***	−0.08***	−0.09***	0.52***	0.53***	0.55***
	(−15.37)	(−7.68)	(−6.00)	(22.38)	(10.86)	(8.86)
PWT 8.0–PWT 6.1	−0.05***	−0.04***	−0.05***	0.05*	0.05	0.055
	(−10.56)	(−4.59)	(−3.56)	(1.91)	(0.94)	(0.77)

Robust t-statistics in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Third, we repeat all estimations using a balanced panel. The purpose of balancing lies in excluding countries for which we have little data with presumably less quality. Assuming that data collection methods are identical or at least similar in a given country over time, using a balanced dataset can produce more reliable estimates and reduce the influence of outliers and measurement error, especially when controlling for country fixed effects. Further, using a balanced sample enables us to investigate if specific country-year observations are driving any of the key results.

Fourth, we then take a step back and consider the set of countries included in our analyses. As noted in Feldkircher and Zeugner (2012), differences in estimation and inference across data revisions of international income data occur not so much because of inherent differences in the methodology of data construction or different data, but the fact that the same set of countries is not used across the datasets. This would imply an apples to oranges comparison of results. To further test the robustness of the results derived in Ram (2009), we re-estimate our regressions, using sets of countries that are consistent between the PWT versions. To get a better sense for how these new datasets are expanding the sample, Table 2 presents the number of countries and observations in common between each of the new datasets and the corresponding PWT 6.1 dataset that Ram (2010)) uses. We see that PWT 7.1 has essentially the same coverage as PWT 6.1 and the PWT 8.0 dataset contains well over 90% of the countries for the time averaged datasets. To determine if it is the methodology of the data construction and not the inclusion/exclusion of countries, we perform regressions using both the PWT 6.1 variables and the alternative datasets' variables on the same set of countries. Specifically, we re-estimate Eqs. (1) to (3) with the additional data sources (PWT 7.1, and PWT 8.0), but keep the same set of sample countries used in the original analysis of Ram (2009) and also the same timeframe from 1960 to 2000.

Finally, we focus on the way our variables are measured. Most importantly, Ram (2009) employs the natural logarithm of government spending as a share of GDP, whereas AW do not use logarithms of government spending or openness. Although the application of the natural logarithm is convenient in macroeconomic variables, especially when comparing a broad spectrum of countries, this methodology is not necessarily superior. Especially in the case of government size and trade openness, which are both measured in percentage terms of GDP, one may well argue for using the initial data. In fact, one usually applies the natural logarithm to variables measured in absolute terms, such as GDP or population size, but not necessarily to scaled variables that are measured in percentages or indices. Notice that out of the incorporated variables here – government size, trade openness, country size (population), and GDP per capita – this argument applies to government size and trade openness, as both variables are initially measured as fractions of GDP. Therefore, we replicate Eqs. (1) through (3) estimating government size and trade openness as pure shares of GDP, absent the natural logarithm.

An overall comparison of our results then allows us to determine whether the results derived by Ram (2009) are driven by one or several of the following aspects: (i) the data source used, (ii) the specific timeframe, (iii) any outliers or countries with few observations, (iv) the countries included in the PWT 6.1 dataset, and (v) the measurement of government size and trade openness.

Beyond the estimation of these models across our different scenarios, we then offer a more detailed analysis on the general relationships displayed in Eqs. (1) to (3). As OLS estimates are by design measuring the conditional mean (under the assumption of correct specification) of the sample, we also focus on conditional quantile estimation for all three models. Focusing on the conditional quantiles of the distribution of the regress and of each model can provide a more detailed evolution of the relationship with the regressors, which the conditional mean may overlook. In this context, we deploy recently developed fixed-effect panel quantile methods (Canay, 2011) to recover the coefficient estimates across vigintiles (quantiles for 0.05, 0.10, 0.15, etc.) for each model.

The panel quantile estimator for the fixed effect framework of Canay (2011) is straightforward to implement in practice. It is a two step estimator, which consists of a first stage standard panel regression where the fixed effects are recovered. These estimated fixed effects are then subtracted from the left hand side variable and this new left hand side variable is then regressed on the covariates in standard quantile fashion. Canay (2011) proves that this two-step estimator provides a consistent estimator for the unknown parameters of the model.

However, one caveat is worth mentioning here. The estimator of Canay (2011) is developed assuming only individual specific heterogeneity. Thus, even though Ram (2009) implements a standard two-way fixed effect framework, our quantile estimates control exclusively for country specific heterogeneity, but not for time specific variation. Nevertheless, we expect the quantile analyses to prove illustrative regarding the behavior of the requisite conditional distributions of interest.

3. Empirical findings

Given that our discussion will revolve around a great number of regression estimates, all tables only display the respective coefficient estimate of interest (the estimated coefficient on *LSIZE* in models (1) and (2) and *LOPEN* in model (3)) along with robust

Table 4
Extension of Ram (2009, Table 1). Dependent variable is the government spending as a percentage of GDP, while all coefficient estimates displayed are for the logarithm of population. All regressions control for the logarithm of GDP per capita.

	OLS			Two-way fixed effects		
	Annual	5-year mean	10-year mean	Annual	5-year mean	10-year mean
1960–2000						
PWT 6.1	−2.21***	−2.33***	−2.47***	5.67***	5.56***	5.54***
	(−19.21)	(−9.29)	(−7.09)	(7.30)	(3.50)	(2.73)
PWT 7.1	−1.90***	−1.89***	−1.88***	−0.95**	−0.31	0.19
	(−32.04)	(−15.07)	(−11.15)	(−2.16)	(−0.40)	(0.20)
PWT 8.0	−0.01***	−0.01***	−0.01***	−0.04***	−0.03*	−0.03
	(−14.68)	(−7.19)	(−5.04)	(−3.46)	(−1.77)	(−1.61)
1960–2010						
PWT 7.1	−1.91***	−1.90***	−1.90***	−2.18***	−1.90**	−1.70
	(−36.02)	(−16.87)	(−12.48)	(−5.28)	(−2.26)	(−1.60)
PWT 8.0	−0.02***	−0.01***	−0.01***	−0.08***	−0.06***	−0.06***
	(−19.58)	(−9.38)	(−6.69)	(−9.61)	(−4.27)	(−3.01)
Balanced 1960–2000						
PWT 6.1	−0.90***	−1.17***	−1.18***	7.34***	7.30***	7.56***
	(−9.18)	(−5.79)	(−4.27)	(9.21)	(4.38)	(3.57)
PWT 7.1	−1.44***	−1.42***	−1.89***	−1.13***	−0.99	0.62
	(−18.74)	(−8.35)	(−11.07)	(−2.66)	(−1.17)	(0.63)
PWT 8.0	−0.01***	−0.01***	−0.01***	−0.05***	−0.05***	−0.02
	(−7.66)	(−4.00)	(−5.56)	(−4.07)	(−2.39)	(−1.22)
Balanced 1960–2010						
PWT 7.1	−1.30***	−1.28***	−1.88***	−2.67***	−2.87***	−1.04
	(−19.51)	(−8.72)	(−12.18)	(−5.99)	(−2.96)	(−0.93)
PWT 8.0	−0.01***	−0.01***	−0.01***	−0.12***	−0.12***	−0.09***
	(−9.46)	(−4.91)	(−7.06)	(−11.21)	(−6.31)	(−4.64)

Robust t-statistics in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Table 5
Replication and extensions of Ram (2009, Table 2). Dependent variable is the logarithm of openness, while all coefficient estimates displayed are for the logarithm of population. All regressions control for the logarithm of GDP per capita.

	OLS			Two-way fixed effects		
	Annual	5-year mean	10-year mean	Annual	5-year mean	10-year mean
1960–2000						
PWT 6.1	−0.21***	−0.20***	−0.20***	0.03*	0.08	0.09
	(−55.70)	(−25.82)	(−19.56)	(1.10)	(1.35)	(1.26)
PWT 7.1	−0.19***	−0.18***	−0.18***	0.01	0.03	0.01
	(−60.09)	(−27.86)	(−20.67)	(0.29)	(0.45)	(0.12)
PWT 8.0	−0.21***	−0.20***	−0.20***	−0.08***	−0.06	−0.07
	(−59.38)	(−27.58)	(−20.22)	(−2.48)	(−1.05)	(−0.94)
1960–2010						
PWT 7.1	−0.16***	−0.16***	−0.16***	0.02	0.04	0.02
	(−58.21)	(−26.97)	(−19.75)	(0.84)	(0.74)	(0.36)
PWT 8.0	−0.18***	−0.18***	−0.17***	−0.08***	−0.07	−0.07
	(−56.15)	(−26.04)	(−19.10)	(−3.40)	(−1.47)	(−1.20)
Balanced 1960–2000						
PWT 6.1	−0.21***	−0.21***	−0.21***	0.04	0.05	0.07
	(−43.36)	(−21.61)	(−15.77)	(1.13)	(0.77)	(0.90)
PWT 7.1	−0.22***	−0.22***	−0.18***	−0.08**	−0.04	0.02
	(−49.92)	(−23.07)	(−20.63)	(−2.05)	(−0.53)	(0.22)
PWT 8.0	−0.24***	−0.24***	−0.21***	−0.10***	−0.06	−0.04
	(−51.47)	(−24.01)	(−20.87)	(−2.85)	(−0.85)	(−0.57)
Balanced 1960–2010						
PWT 7.1	−0.20***	−0.20***	−0.16***	−0.07***	−0.05	0.04
	(−46.94)	(−21.52)	(−19.94)	(−2.43)	(−0.87)	(0.54)
PWT 8.0	−0.21***	−0.21***	−0.18***	−0.12***	−0.09	−0.06
	(−48.32)	(−22.42)	(−19.56)	(−3.81)	(−1.38)	(−0.98)
Using consistent countries, 1960–2000						
PWT 6.1–PWT 7.1	−0.21***	−0.20***	−0.21***	0.37***	0.39***	0.38***
	(−55.49)	(−25.78)	(−19.69)	(21.04)	(11.14)	(8.53)
PWT 7.1–PWT 6.1	−0.21***	−0.20***	−0.20***	0.20***	0.21***	0.21***
	(−56.58)	(−25.97)	(−19.62)	(10.90)	(5.62)	(4.17)
PWT 6.1–PWT 8.0	−0.21***	−0.21***	−0.21***	0.37***	0.39***	0.39***
	(−55.26)	(−25.57)	(−19.35)	(20.55)	(10.64)	(8.13)
PWT 8.0–PWT 6.1	−0.22***	−0.21***	−0.21***	0.31***	0.32***	0.31***
	(−55.47)	(−25.43)	(−18.98)	(16.70)	(8.32)	(6.10)

Robust t-statistics in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

t-statistics in parentheses underneath the estimates.[5] For the sake of replicability we present the robust t-statistics as in Ram (2009), however, reporting t-statistics is risible given that common empirical practice is to present *p*-values and focus discussion around the economic meaning and the statistical precision (usually through confidence intervals) of the estimates. By its nature a *t*-statistic is relatively nebulous when characterizing statistical precision, relating solely to a specific null hypothesis.

Following Ram (2009) structure, the first three columns in all tables show the results from pooled regressions and the final three columns include country and time fixed effects. Rows 2 to 14 of Tables 3–8 look at the results for each model of Ram (2009). Tables 3 and 4 show all our results for the potential effects of country size on government size. Tables 5 and 6 then move to the second part of AW's theory, suggesting that smaller countries are more open to international trade, given their limited domestic markets. Finally, Tables 7 and 8 analyze whether more open economies indeed have bigger governments, the argument proposed by Rodrik (1998) and solidified by Ram (2009). It is worth mentioning that the main contribution of Ram (2009) is that all three results need to hold to conclusively refute the findings of AW, namely that country size is not driving the link between trade openness and government size, while supporting the insights of Rodrik (1998).

Figs. 1 to 3, in addition, present quantile coefficient plots for the main coefficient of interest from each model. The solid black circles represent the estimates from panel quantile estimation, while the hollow, red circles are the standard, cross sectional quantile estimates from pooled regressions (using annual data). The lines stemming from the circles display the 95% confidence intervals for each vigintile. Finally, Table 9 combines our findings and presents results from regressing government size on both country size and trade openness in a single regression framework.

[5] Robust *t*-statistics are calculated using the method of White (1980) allowing for clustering over time, general heteroskedasticity across countries, but no serial correlation, using the HC_0 form.

Table 6
Extension of Ram (2009, Table 2). Dependent variable is openness to trade as a percentage of GDP, while all coefficient estimates displayed are for the logarithm of population. All regressions control for the logarithm of GDP per capita.

	OLS			Two-way fixed effects		
	Annual	5-year mean	10-year mean	Annual	5-year mean	10-year mean
1960–2000						
PWT 6.1	−12.56***	−12.62***	−12.71***	1.32**	4.30	4.60
	(−50.41)	(−23.97)	(−18.47)	(0.55)	(0.94)	(0.79)
PWT 7.1	−11.13***	−10.95***	−10.82***	3.30	5.55	3.87
	(−55.12)	(−25.77)	(−19.04)	(1.26)	(1.24)	(0.72)
PWT 8.0	−0.13***	−0.13***	−0.12***	−0.02	−0.01	−0.03
	(−54.15)	(−25.30)	(−18.58)	(−0.72)	(−0.32)	(−0.61)
1960–2010						
PWT 7.1	−10.09***	−9.96***	−9.79***	1.13	2.81	1.56
	(−53.65)	(−25.05)	(−18.36)	(0.59)	(0.80)	(0.35)
PWT 8.0	−0.11***	−0.11***	−0.11***	−0.05***	−0.04	−0.05
	(−51.97)	(−24.20)	(−17.78)	(−2.56)	(−1.17)	(−1.17)
Balanced 1960–2000						
PWT 6.1	−11.01***	−12.20***	−12.26***	4.01*	3.13	4.26
	(−36.19)	(−18.02)	(−13.12)	(1.70)	(0.64)	(0.70)
PWT 7.1	−12.37***	−12.39***	−10.89***	3.99	6.34	4.42
	(−42.82)	(−19.73)	(−18.83)	(1.64)	(1.25)	(0.80)
PWT 8.0	−0.13***	−0.14***	−0.13***	0.02	0.05	0
	(−42.46)	(−19.68)	(−18.87)	(0.63)	(0.87)	(0.00)
Balanced 1960–2010						
PWT 7.1	−11.78***	−11.78***	−9.97***	0.15	1.83	3.09
	(−41.33)	(−18.98)	(−18.31)	(0.08)	(0.45)	(0.66)
PWT 8.0	−0.13***	−0.13***	−0.12***	−0.02	0.01	−0.02
	(−41.38)	(−19.10)	(−18.03)	(−0.89)	(0.18)	(−0.38)

Robust t-statistics in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

3.1. Do smaller countries have bigger governments?

Starting with Table 3, the results in the first row reproduce the key finding of Ram (2009) from estimating model (1). Country size produces a negative and statistically precise impact on government size in the pooled data setting. Both AW and Ram (2009) produce tight estimates of this impact in the pooled setting. However, the estimated impact of country size is reversed when accounting for country and time specific heterogeneity — the effect of country size on the size of government is now statistically precise and positive. Moreover, the magnitude of the estimated effect is on the order of about 5 times larger than in the pooled setting. This finding provides clear evidence against AW's theoretical underpinnings as both the results of models (1) and (2) are required to cast doubt on Rodrik (1998) theory (model (3)).

Rows 2 and 3 then use the same timeframe (1960–2000), but employ data from the two recent version of the PWT (7.1 and 8.0). Although the results from pooled OLS regressions are replicating row 1 in terms of statistical importance, magnitudes of the derived coefficients vary substantially between −0.05 and −0.12. Including two-way fixed effects then completely alters Ram's (2009) key insight: country size returns as a coefficient that is not distinguishable from zero in its effect on government size when using PWT 7.1 data, but then produces a negative coefficient with data from PWT 8.0, as initially suggested by AW. Throughout all three datasets, the coefficient for using annual observations ranges from −0.17 to +0.48, which means that across the three datasets, the estimated results for Eq. (1) allow for all possible conclusions. Notice that country size is a negative and statistically significant determinant (on the one percent level) of government size when using annual data and PWT 8.0 data. Thus, depending on the dataset chosen, the derived conclusions differ dramatically, both in terms of sign and statistical relevance.

Rows 4 and 5 then extend the timeframe towards including 2001–2010. Given the initial data limitations of PWT 6.1, only the PWT 7.1 and 8.0 versions incorporate these years. The results from these alternative datasets provide additional support for AW's insights, further casting doubt on the generality of Ram's (2009) results. It appears as if smaller countries are indeed marked by bigger governments. Notice that the statistical importance of this finding is unquestioned for PWT 8.0 data, whereas for the PWT 7.1 data, the result only emerges with strong statistical evidence for annual data. Nevertheless, these results are markedly different from the coefficients displayed in row (1). This means that not only does it matter which dataset we choose, but also the selected (or available) timeframe can influence the derived relationship between country size and government size. Another interpretation of comparing rows 4 and 5 to rows 2 and 3 could be that the link between these two variables changes over time. Applied to our results here, this would mean that smaller countries exhibit bigger governments especially in recent years, but to a lesser degree before the 2000s.

Next, we focus on the countries included across the respective datasets. Rows 6 through 8 show results choosing the initial timeframe from 1960 to 2000, but balancing the sample for all datasets. This means that only countries are included for which all

Table 7
Replication and extensions of Ram (2009, Table 3). Dependent variable is the logarithm of government spending, while all coefficient estimates displayed are for the logarithm of openness. All regressions control for the logarithm of GDP per capita.

	OLS			Two-way fixed effects		
	Annual	5-year mean	10-year mean	Annual	5-year mean	10-year mean
1960–2000						
PWT 6.1	0.24***	0.27***	0.28***	0.08***	0.11***	0.16***
	(17.08)	(8.46)	(6.02)	(4.51)	(2.63)	(2.73)
PWT 7.1	0.22***	0.24***	0.25***	0.03***	0.04*	0.06*
	(17.08)	(8.45)	(6.17)	(2.54)	(1.83)	(1.86)
PWT 8.0	0.15***	0.16***	0.16***	−0.01**	−0.01	0.01
	(12.64)	(6.19)	(4.68)	(−0.48)	(−0.23)	(0.22)
1960–2010						
PWT 7.1	0.20***	0.22***	0.23***	0.09***	0.10***	0.12***
	(18.03)	(8.84)	(6.58)	(7.57)	(4.50)	(3.76)
PWT 8.0	0.14***	0.14***	0.15***	0.03*	0.03	0.03
	(12.95)	(6.28)	(4.68)	(1.67)	(0.72)	(0.60)
Balanced 1960−2000						
PWT 6.1	0.13***	0.16***	0.17***	0.09***	0.12***	0.16***
	(8.38)	(5.10)	(3.74)	(4.56)	(2.71)	(2.58)
PWT 7.1	0.14***	0.15***	0.26***	−0.01	−0.00	0.05
	(9.81)	(4.64)	(6.29)	(−0.69)	(−0.08)	(1.65)
PWT 8.0	0.08***	0.08***	0.13***	−0.05***	−0.03	−0.00
	(5.81)	(2.53)	(3.75)	(−2.40)	(−0.75)	(−0.03)
Balanced 1960–2010						
PWT 7.1	0.13***	0.14***	0.24***	0.07***	0.09***	0.10***
	(10.03)	(4.73)	(6.54)	(4.45)	(2.74)	(3.24)
PWT 8.0	0.05***	0.04	0.12***	0.03	0.05	0.03
	(4.10)	(1.59)	(3.66)	(1.40)	(1.23)	(0.63)
Using consistent countries, 1960−2000						
PWT 6.1–PWT 7.1	0.25***	0.27***	0.27***	0.21***	0.28***	0.38***
	(17.14)	(8.48)	(5.90)	(10.45)	(6.31)	(5.70)
PWT 7.1–PWT 6.1	0.20***	0.21***	0.23***	0.02	0.05*	0.08*
	(13.87)	(6.90)	(5.11)	(1.38)	(1.70)	(1.82)
PWT 6.1–PWT 8.0	0.23***	0.26***	0.26***	0.22***	0.30***	0.39***
	(15.47)	(7.76)	(5.47)	(10.67)	(6.43)	(5.68)
PWT 8.0–PWT 6.1	0.18***	0.19***	0.18***	0.02	0.05	0.08
	(13.57)	(6.69)	(4.61)	(0.78)	(0.96)	(1.13)

Robust t-statistics in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

observations are available. Comparing the derived coefficient estimates to rows 1 through 3 further solidifies the previous conclusions: we can deduct a positive, a negative, or a statistically insignificant coefficient on country size, depending on which dataset we choose. The same conclusion can be drawn when balancing the entire available timeframe from 1960 to 2010 (rows 9 and 10) and comparing these estimates to rows 4 and 5. Overall, balancing seems to strengthen the derived relationships by marginally increasing the initial coefficients.

Finally, rows 11 through 14 re-estimate Eq. (1) for the timeframe 1960 to 2000, but considering a consistent set of sample countries. Every row uses only the observations, which are common across both mentioned datasets. For instance, row 11 (PWT 6.1–PWT 7.1) uses the PWT 6.1 data, but only includes the country-year observations, which can also be found in the PWT 7.1 sample. Similarly, row 12 (PWT 7.1–PWT 6.1) employs that same set of country-year observations using the PWT 7.1 data.

Not surprisingly, the pooled estimations produce a consistently negative elasticity, which is statistically significant and displays magnitudes consistent with earlier estimates. We observe this result throughout all of the data settings in our study, yet the disparities are setting in when controlling for country and time specific effects. These estimated models produce telling insights: all estimated models now produce a positive coefficient for country size, with 10 out of 12 leading to a coefficient estimate that is statistically significant on conventional levels. This means that the choice of sample countries is absolutely crucial when estimating the relationship between population size and the extent of the country's government. In particular, the countries included in PWT 6.1 seem to favor a positive relationship between country size and government size. In additional exercises (available upon request), we balanced the datasets in rows 11 through 14 and uncovered similar results.

As a final exercise related to the link between country size and the extent of the public sector, we now consider the measurement of government size. Table 4 replicates the main results from Table 3 using the pure percentage of government size in GDP, without applying the natural logarithm. In fact, this is the initial form of measurement employed by AW. A close look, however, reveals that there are only few notable differences between the two estimation techniques when it comes to suggested significance levels. Note

Table 8
Extension of Ram (2009, Table 3). Dependent variable is government spending as a percentage of GDP, while all coefficient estimates displayed are for openness to trade as a percentage of GDP. All regressions control for the logarithm of GDP per capita.

	OLS			Two-way fixed effects		
	Annual	5-year mean	10-year mean	Annual	5-year mean	10-year mean
1960–2000						
PWT 6.1	0.07***	0.08***	0.08***	0.05***	0.07***	0.09***
	(11.67)	(5.61)	(4.01)	(7.37)	(4.50)	(4.28)
PWT 7.1	0.04***	0.04***	0.04***	−0.00	0.00	0.01
	(14.13)	(6.90)	(4.86)	(−0.28)	(0.32)	(1.53)
PWT 8.0	0.03***	0.03***	0.03***	−0.01	−0.02*	−0.01
	(7.01)	(3.35)	(2.47)	(−1.27)	(−1.74)	(−0.75)
1960–2010						
PWT 7.1	0.03***	0.04***	0.04***	0.01**	0.01*	0.02**
	(14.40)	(6.94)	(5.10)	(2.21)	(1.86)	(2.27)
PWT 8.0	0.02***	0.02***	0.02***	−0.01*	−0.02***	−0.02*
	(7.00)	(3.28)	(2.37)	(−1.71)	(−2.37)	(−1.80)
Balanced 1960–2000						
PWT 6.1	0.06***	0.04***	0.04***	0.06***	0.08***	0.10***
	(9.67)	(3.78)	(2.72)	(7.67)	(4.50)	(3.82)
PWT 7.1	0.02***	0.02***	0.04***	−0.02***	−0.03***	0.01
	(8.65)	(4.20)	(5.09)	(−4.88)	(−2.54)	(1.36)
PWT 8.0	0.00	0.00	0.02*	−0.04***	−0.05***	−0.02
	(1.26)	(0.40)	(1.94)	(−4.58)	(−2.46)	(−1.31)
Balanced 1960–2010						
PWT 7.1	0.02***	0.02***	0.04***	−0.01*	−0.01	0.01
	(8.89)	(4.25)	(5.16)	(−1.85)	(−0.94)	(1.65)
PWT 8.0	0.00	−0.00	0.02*	−0.02***	−0.02	−0.02*
	(−0.12)	(−0.46)	(1.87)	(−2.44)	(−1.17)	(−1.67)

Robust t-statistics in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

that the quantitative interpretation changes when considering percentages. Overall, using the pure fraction of government spending in GDP tends to favor AW's story of smaller countries exhibiting bigger governments, whereas the logarithmic transformation supports Ram's (2009) conclusion.

In summary, Ram's (2009) conclusions regarding the link between country size and government size are highly dependent on the dataset employed, the timeframe chosen, and the countries included. Not only do we find sharp differences in magnitudes, but rather

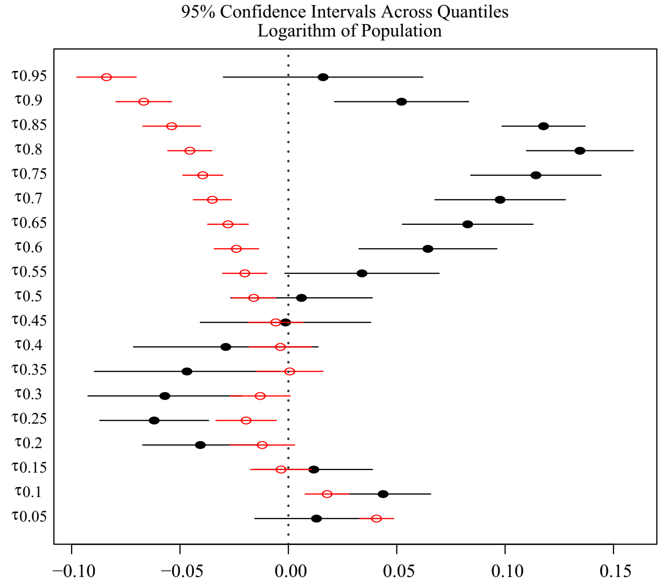

Fig. 1. Vigintile estimates for Model 1 of Ram (2009): The effect of country size on government size.

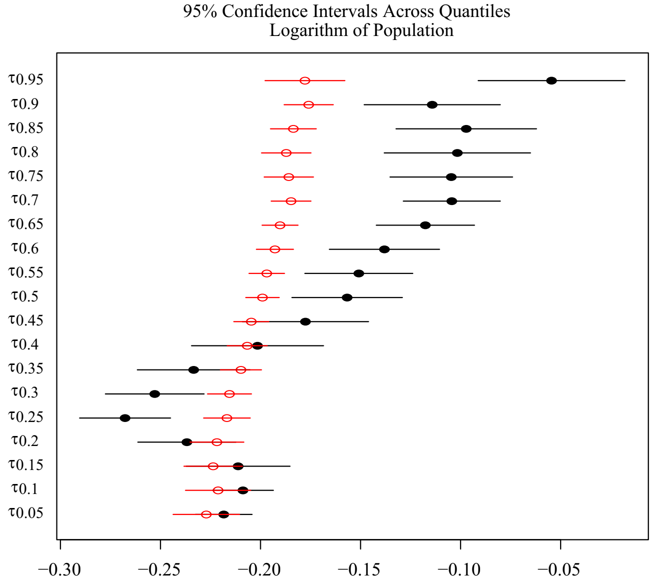

Fig. 2. Vigintile estimates for Model 2 of Ram (2009): The effect of country size on trade openness.

a move from one end of the statistical spectrum to the other, depending on these alternative scenarios. This is troubling given that the switch from a negative and statistically precise estimated elasticity to one that is positive and statistically precise is one of the key results to lend support to the theoretical model of Rodrik (1998) by overturning one of the channels through which a country's openness operates on government size, namely country size. A natural question emerging from these conclusions is then why using the PWT 6.1 data produces these results.

Some insight can be gleaned from Fig. 1, visualizing the vigintile estimates on the estimated elasticity in model (1) using the balanced sample of countries in PWT 6.1 over the 1960–2000 timeframe. Fig. 1 shows that Ram's (2009) result holds for the 60th to the 90th vigintiles ($\tau_{0.60} - \tau_{0.90}$). This suggests that it is this segment of the distribution that is driving the conditional mean results witnessed in Table 3, row 1. The cross-sectional 5th and 10th vigintile estimates are positive while the 15th, 20th, 30th, 35th, 40th and 45th vigintile estimates are negative, but statistically imprecise. These vigintile estimates also caution against focusing exclusively on the conditional mean. Even though the balanced PWT 6.1 sample over the 1960–2000 period supports Ram's (2009) insights, were

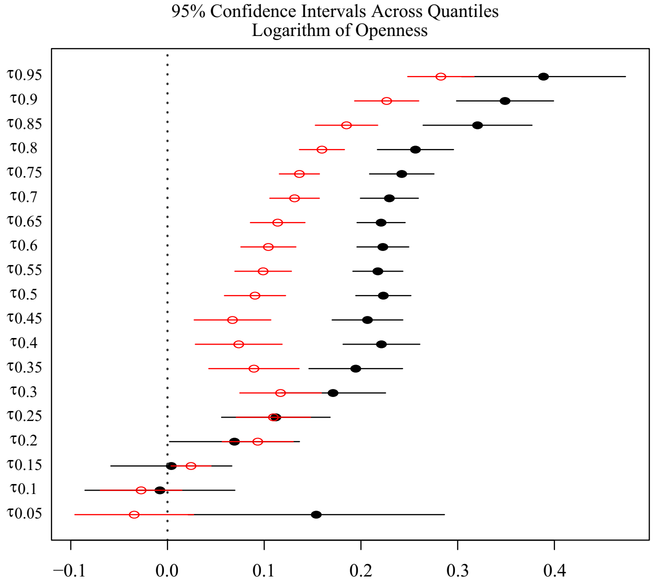

Fig. 3. Vigintile estimates for Model 3 of Ram (2009): The effect of trade openness on government size.

Table 9
Extension of Ram (2009, Table 3). Dependent variable is the logarithm of government spending, while coefficient estimates displayed are for the logarithm of openness (top row) and the logarithm of population (bottom row). All regressions control for the logarithm of GDP per capita.

	OLS			Two-way fixed effects		
	Annual	5-year mean	10-year mean	Annual	5-year mean	10-year mean
1960–2000						
PWT 6.1	0.15***	0.17***	0.15***	0.08***	0.10***	0.15***
	(8.70)	(4.30)	(2.60)	(4.37)	(2.44)	(2.56)
	−0.05***	−0.05***	−0.06***	0.47***	0.43***	0.41***
	(−8.33)	(−3.86)	(−3.30)	(10.77)	(4.85)	(3.54)
PWT 7.1	−0.01	−0.00	−0.02	0.03***	0.05*	0.06*
	(−0.91)	(−0.10)	(−0.40)	(2.57)	(1.83)	(1.90)
	−0.13***	−0.13***	−0.13***	−0.00	0.02	0.04
	(−27.71)	(−12.76)	(−9.34)	(−0.08)	(0.32)	(0.47)
PWT 8.0	0.09***	0.10***	0.12**	−0.01	−0.01	0.01
	(5.13)	(2.78)	(2.31)	(−0.62)	(−0.28)	(0.18)
	−0.04***	−0.03***	−0.02	−0.17***	−0.12	−0.10
	(−6.59)	(−2.44)	(−1.34)	(−4.77)	(−1.64)	(−1.11)
1960–2010						
PWT 7.1	−0.00	0.00	0.00	0.09***	0.11***	0.12***
	(−0.26)	(0.14)	(0.08)	(7.50)	(4.46)	(3.72)
	−0.13***	−0.13***	−0.13***	−0.07***	−0.07	−0.07
	(−33.39)	(−15.47)	(−11.28)	(−2.76)	(−1.28)	(−1.00)
PWT 8.0	0.05***	0.06**	0.07*	0.02	0.02	0.02
	(3.71)	(2.12)	(1.72)	(1.21)	(0.54)	(0.45)
	−0.05***	−0.05***	−0.04***	−0.37***	−0.29***	−0.26***
	(−12.22)	(−4.85)	(−3.23)	(−11.57)	(−4.38)	(−3.12)
Balanced 1960–2000						
PWT 6.1	0.13***	0.14***	0.15***	0.09***	0.12***	0.15***
	(6.57)	(3.47)	(2.51)	(4.38)	(2.61)	(2.46)
	−0.00	−0.01	−0.01	0.55***	0.52***	0.54***
	(−0.44)	(−0.93)	(−0.69)	(11.45)	(5.44)	(4.34)
PWT 7.1	−0.03	−0.02	−0.01	−0.01	−0.00	0.06
	(−1.51)	(−0.58)	(−0.29)	(−0.69)	(−0.07)	(1.68)
	−0.12***	−0.12***	−0.13***	0.01	0.01*	0.08
	(−16.06)	(−7.04)	(−9.24)	(0.35)	(0.17)	(1.03)
PWT 8.0	0.03*	0.03	0.05	−0.05***	−0.04	−0.00
	(1.74)	(0.69)	(0.95)	(−2.75)	(−0.84)	(−0.05)
	−0.03***	−0.03*	−0.04***	−0.30***	−0.25***	−0.05
	(−4.14)	(−1.81)	(−2.37)	(−7.67)	(−2.96)	(−0.53)
Balanced 1960–2010						
PWt 7.1	−0.01	−0.00	0	0.07***	0.09***	0.11***
	(−0.33)	(−0.06)	(−0.00)	(4.22)	(2.61)	(3.21)
	−0.10***	−0.10***	−0.13***	−0.04	−0.07	−0.01
	(−16.74)	(−7.43)	(−10.91)	(−1.46)	(−1.04)	(−0.14)
PWT 8.0	−0.01	−0.02	0.01	0.01	0.04	0.02
	(−0.61)	(−0.57)	(0.25)	(0.60)	(0.93)	(0.48)
	−0.04***	−0.04***	−0.06***	−0.55***	−0.54***	−0.36***
	(−7.46)	(−3.46)	(−4.18)	(−14.26)	(−6.21)	(−3.86)

Robust t-statistics in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

we to conduct the same exercise at the median (which for a symmetric distribution is equivalent to the mean) we would not find the same results. Even in this case our estimated elasticity in the pooled model is negative and statistically significant, but the fixed effect panel estimate is positive and statistically insignificant.

The main take away from Fig. 1 consists in recognizing that there exist substantial differences throughout the conditional distribution of government spending and strictly placing our attention on the conditional mean may not be appropriate regarding adjudication across empirical models. More generally, we see that the result of AW holds in the lower portion of the distribution of government size, but recedes as we move to the upper portion of the distribution, conditional on country size and income. That is, it could be that countries with relatively smaller governments are impacted by the size of the populace, while countries that are relatively larger are no longer affected — a simple marginal product type argument.

After the in-depth analysis of the link between country size and government size, we now move to Eq. (2), considering whether the size of a country is indeed related to its degree of trade openness, as suggested by AW.

3.2. Are smaller countries more open to international trade?

The second argument brought forward by Alesina and Wacziarg (1998) suggests a natural relationship between the size of a country and its openness to international trade. As smaller countries also exhibit smaller domestic markets, their tendency to look for international trading partners may naturally be emphasized. Table 5 displays the estimated elasticities between trade openness and country size across all data settings for Eq. (2).

Row (1) replicates Ram's (2009) estimates of the relationship between country size and trade openness. Once again the inclusion of two-way fixed effects leads to a substantial change in results. The negative and statistically precise relationship that exists in the pooled sample disappears in the estimation of the model in the fixed effects framework — a finding that is consistent when using annual values or averaging over 5 and 10 year periods. The pooled estimated elasticity is consistent with −0.2, but when accounting for the panel structure the estimates are positive, with elasticity levels ranging from 0.03 to 0.09. This is consistent with the unobserved time and country specific variation producing much of the impact detected in the pooled estimates. Noting both the sign change of the estimated elasticity as well as the relatively wide confidence intervals, the main conclusion of Ram (2009) is that an economically or statistically strong relationship between country size and openness to trade does not exist. Coupled with the findings from row (1) of Table 3, the estimation of AW's models accounting for unobserved country and time specific effects casts doubt on the omitted "country-size" effect that might drive the link between trade openness and government size.

Similar to Table 3, we now conduct different regressions, trying to pin down whether the derived result changes across datasets, timeframes, sample countries, and measurement technique. These extensions once again reveal differences from Ram's (2009) findings, although the contrast is not as glaring. Throughout rows 2 to 10, we have coefficient estimates that almost entirely are not different from zero in a statistical sense. The only exceptions are provided by the PWT 8.0 data (and the PWT 7.1 data when balancing the sample) when using annual observations — in this case, we find evidence for AW's claim that smaller countries indeed exhibit more openness to international trade.

However, these results change dramatically when we streamline the countries included in each dataset (rows 11 to 14). In fact, Ram's (2009) results are substantially more consistent when sticking to the initial set of observations, regardless of dataset. The estimated coefficient on the logarithm of population goes from negative and significant to positive and significant when migrating from the pooled setting to the model that accounts for unobserved country and time specific heterogeneity. In fact, all of the setups incorporating two-way fixed effects produce positive coefficient estimates for *LSIZE* and all of them show significance at the 1% level with magnitudes fluctuating between 0.2 and 0.39. This suggests that an increase in the population by 10% would correspond to an increase in trade openness by 2 to 3.9%. Comparing these findings to the extensions along the lines of timeframes in rows 4, 5, 9, and 10 then once again shows that the choice of sample countries is crucial.

Beyond datasets, timeframes, and selected countries, Table 6 now evaluates whether the initial findings are robust to using trade as a pure fraction of GDP, rather than taking the natural logarithm of that variable. Replicating the first 10 rows of Table 5, we find very little evidence that the size of a country is related to trade openness. Only two out of 30 potential scenarios produce statistically significant coefficients, which are inconclusive as they point in opposite directions. The form of measurement does seem to matter here, more so than for Eq. (1).

Overall, the results across different data settings are inconclusive as to whether country size is driving trade openness. Certainly, Ram's (2009) rejection of AW's theory of a negative link seems to be an artifact of using the PWT 6.1 data and its sample countries, but the way of measuring trade openness also plays a substantial role. We do find some evidence for a negative relationship (albeit weak) when using PWT 8.0. However, these clues are not nearly as strong as in the previously discussed link between country size and government size.

Finally, Fig. 2 takes a closer look at the link between trade openness and country size throughout the distribution. Using country fixed effects reveals that the negative link is especially prevalent for relatively closed economies. As countries are becoming more open to trade, the coefficient estimate moves towards zero and eventually settles around −0.1. The pooled quantile estimates reveal a near constant elasticity throughout the distribution. Unlike the conditional mean results presented in Table 5, our quantile estimates reveal a negative elasticity (albeit ignoring time effects). Thus, it appears that unobserved time heterogeneity plays a key role in Ram's (2009) conditional mean findings for the link between country size and trade openness.

3.3. Do open economies have bigger governments?

After considering the role of country size in explaining the size of the public sector, we now move to the suggested link between trade openness and government size. Row (1) of Table 7 replicates Ram's (2009) estimation of this relationship (Rodrik (1998) model), displaying the derived coefficient for trade openness. Here, the estimates from the pooled regressions are sensitive to the inclusion of two-way fixed effects in magnitude only. The estimated elasticity in these pooled models comes out to be roughly 0.25. When accounting for unobserved country and time specific heterogeneity, elasticities are statistically precise with the same positive sign, whereas the magnitude of these estimated elasticities is cut by about a half to two thirds. Of the three regression models, only model (3) produces estimates that economically have the same meaning across pooled and panel models as with the initial estimates using PWT 6.1.

Re-estimating the initial regression with PWT 7.1 data then produces similar conclusions, although the estimated coefficients in the fixed effect framework decrease marginally. Applying the PWT 8.0 data leads to coefficients that are statistically not distinguishable from zero and even turn marginally negative. In fact, this impression from using the PWT 8.0 data remains for the majority of estimates displayed in Table 7, as the coefficient on trade openness remains statistically insignificant when extending the timeframe, balancing the sample, or restricting the sample to countries available in PWT 6.1 only. Whereas the results for using the PWT 6.1 data are the strongest throughout this table, the PWT 7.1 data also supports a meaningful and positive relationship between trade openness and government size. For this relationship, it seems as if the choice of dataset is crucial. To see this, consider the final 4 rows of Table 7: using consistent countries, we find a strong positive link when employing the PWT 6.1 data, weak evidence from using the PWT 7.1 data, and no meaningful relationship for using the PWT 8.0 data.

Addressing the form of measurement, Table 8 replicates the possible combinations of using different datasets, timeframes, and sample countries when using the raw shares of trade openness and government spending of GDP. As before, the PWT 6.1 data is the most forceful supporter of more open countries selecting bigger governments, as famously advocated by Rodrik (1998) and later supported by Ram (2009). The PWT 7.1 data produces mixed results, whereas the PWT 8.0 data supports the idea that more open economies have smaller governments — similar to the result found in Benarroch and Pandey (2012). Overall, the PWT 8.0 data produces a negative coefficient estimate across all of the data settings in Table 8 with 8 out of 12 estimated coefficients returning coefficients that are statistically significant on conventional levels with magnitudes in line with past research.

Thus, no clear statistical evidence emerges that there is a positive elasticity between openness to trade and government size and the PWT 8.0 data even leads us to believe that the suggested link is negative. In this third equation, the question of how we measure the respective variables becomes relevant, more so than in the equations assessing the intermediating role of country size. Once again, the proposed relationship is far from being resolved, as results are fragile for various extensions and robustness checks.

Finally, Fig. 3 takes a closer look at this link across different levels of government size. It is interesting to see that for the vigintiles greater than 0.2 the results of Ram (2009) are consistent. As governments get bigger, the positive link between trade openness and government size becomes stronger. One interpretation of this finding is that when governments are very small the institutions for accommodating greater demands for an insurance towards international risk (the channel suggested by Rodrik, 1998) are not in place yet. Thus, the option of strengthening the public safety net may only become available if the respective institutions are in place already.

3.4. Combining findings

To conclude our analysis of these macroeconomic relationships, we now test the general result from our baseline analysis in a horse race as to which theory remains more robust: Which of the two factors – trade openness or country size – dominates in explaining government size? Since we do not find a persistent link between country size and openness itself, we now feel confident to test both theories in a single equation. If indeed openness influences government spending only through its connection to country size, as AW point out, then including country size into the model should wipe away any effect of openness. Table 9 displays the estimates across our various models and data settings. As before, all regressions control for the logarithm of GDP per capita.

First, consider the pooled OLS results throughout all data settings displayed in Table 9. Without accounting for fixed effects, the individual relationships in the previous tables have generally been clear, supporting both theories of AW and Ram (2009). However, this is not the case when including both trade openness and country size in estimating government size. Here, the coefficient estimates are relatively scattered, at least for the coefficient associated with trade openness; especially for PWT 7.1 data, which does not support a meaningful link between openness to international trade and the extent of the public sector. In fact, the coefficient estimates are mostly negative when employing PWT 7.1 data. Nevertheless, incorporating fixed effects has been shown to be important in these macroeconomic analyses, so we do not wish to overemphasize these results from pooled regressions. However, once country- and time-specific unobservables are incorporated, the picture does not become much clearer.

In the initial PWT 6.1 setting (row 1) we find strong support for the openness theory, as the positive and statistically significant coefficient estimate remains, yet country size is suggested to have a positive and significant effect on government size once fixed effects are included. Notice that we find the same switch in signs when moving from the pooled sample to accounting for fixed effects, replicating Ram's (2009) initial findings. These estimates run contrary to AW's theory and in fact leave the opposite conclusion — smaller countries should have smaller governments in relative terms.

Now consider the estimated coefficients when using the alternative, more recent datasets for the same timeframe. First, the PWT 7.1 data supports a positive link between trade openness and government size, but there is no clear connection between country size and government size with signs switching and none of the estimates are statistically significant at conventional levels. Moving to the PWT 8.0 data further changes the derived conclusions: Now, we do find some evidence for smaller countries choosing bigger governments, especially when using annual observations. Trade openness, on the other hand, seems unrelated to government size. Notice that the estimates related to trade openness exhibit robust *t*-values under one, meaning that we are far from levels that are usually accepted as statistically meaningful.

As a next step, we add ten years to the samples in rows 4 and 5. Differences between the two most recent PWT versions could hardly be any bigger, as the PWT 7.1 firmly supports Rodrik (1998) and Ram (2009), whereas the PWT 8.0 data lends substantial support to AW. Notice that we find some evidence for the country size theory in the PWT 7.1 dataset as well. Balancing the data produces a similar picture: depending on which dataset we prefer, we would either conclude that trade openness or country size better predicts government size. In fact, picking a variable that is most consistent in predicting government size proves to be difficult, as country size emerges with a negative coefficient estimate in 19 out of 30 regressions with 10 of them showing statistical relevance. Trade openness, on the other hand, emerges as a positive predictor of government size 17 times with 12 of them showing importance in a statistical sense.

Table 9 is somewhat representative for the overall findings presented in this article. In the end, none of the suggested variables, neither trade openness nor country size, receive strong and robust support in explaining government size. In general, Ram's (2009) findings seem to be an artifact of the dataset available at that time (PWT 6.1), the timeframe available (1960 to 2000), and the specific countries included in that sample. Were we to favor the PWT 6.1 data, then we would conclude trade openness to be an important predictor of government size and the alternative explanation of country size influencing both variables would be discarded. However, using the PWT 8.0 data would lead one to the opposite conclusion. Somewhat in the middle of both extremes lies the PWT 7.1 dataset, where the derived results are even more sensitive to the selected timeframe and countries.

4. Conclusions

This paper casts significant doubt on two conclusions, namely that (i) more open economies exhibit bigger governments and that (ii) country size is playing an insignificant role in that relationship, as suggested by Ram (2009). More specifically, Ram's (2009) results lend support to the government size model of Rodrik (1998) over that of Alesina and Wacziarg (1998) once unobserved country and time specific heterogeneity are accounted for. After successfully replicating the main empirical findings of Ram (2009) we extend his analysis along several dimensions. Specifically, we test whether this result is robust to the choice of (i) dataset, (ii) timeframe, (iii) sample countries, and (iv) the measurement of the respective variables.

Overall, our results suggest that all four of these attributes matter in the discussed relationship. In particular, the PWT 8.0 dataset generally leads to the opposite conclusion, after which country size may be a more important predictor of government size than trade openness. Thus, Ram's, (2009) findings appear to be an artifact of the dataset selected, in that the countries and years included in that dataset favor Rodrik's (1998) results and dismiss Alesina and Wacziarg's 1998 conclusions. Our findings cannot conclusively dismiss the importance of country size in this relationship, and neither can we provide robust support towards the openness theory. Our conclusion from these findings leans towards a cautious interpretation of Ram's (2009) results.

This allows for two basic conclusions. First, the relationship between trade openness and government size is far from being explained. We cannot exclude the possibility of country size playing a decisive intermediating role. And secondly, the discrepancies across major sources for macroeconomic data can not only be found in results pertaining to economic growth, but also when focusing on government size and trade openness.

Assuming that newer data is better data does not help in this context, since the results from using recent data sources (PWT 7.1, released in November 2012, versus PWT 8.0 data, released in June 2013) differ substantially. Differences related to the methodology of calculating variables are discussed in more detail in Feenstra et al. (2011) and the PWT 8.0 user guide. In addition, a relatively small change in sample size, either along the country or the time dimension, can change seemingly robust coefficient estimates, not only in their magnitude, but even in their suggested sign and statistical significance. In fact, our results show that the estimated elasticity of country size can turn from positive and significant to negative and significant when explaining government size. Finally, more broadly, our results provide a cautionary example for researchers, suggesting care when estimating relationships using cross-country macroeconomic data and advising to compare results from using different data sources.

References

Afonso, A., Furceri, D., 2010. Government size, composition, volatility and economic growth. Eur. J. Polit. Econ. 26 (4), 517–532.
Akitoby, B., Clements, B., Gupta, S., Inchauste, G., 2006. Public spending, voracity, and Wagner's law in developing countries. Eur. J. Polit. Econ. 22 (4), 908–924.
Alcalá, F., Ciccone, A., 2004. Trade and productivity. Q. J. Econ. 2, 613–646.
Alesina, A., Wacziarg, R., 1998. Openness, country size and government. J. Public Econ. 69, 305–321.
Atherton, P., Appleton, S., Bleaney, M., 2011. Growth regressions and data revisions in Penn World Tables. J. Econ. Stud. 38 (3), 301–312.
Barro, R.J., 2001. Human capital and growth. Am. Econ. Rev. 91 (2), 12–17.
Barro, R.J., Lee, J.-W., 1994. Sources of economic growth. Carnegie–Rochester Conference Series on Public Policy vol. 40. Elsevier, pp. 1–46.
Benarroch, M., Pandey, M., 2008. Trade openness and government size. Econ. Lett. 3, 157–159.
Benarroch, M., Pandey, M., 2012. The relationship between trade openness and government size: does disaggregating government expenditure matter? J. Macroecon. 34 (1), 239–252.
Bergh, A., Henrekson, M., 2011. Government size and growth: a survey and interpretation of the evidence. J. Econ. Surv. 25 (5), 872–897.
Breton, T.R., 2012. Penn World Table 7.0: are the data flawed? Econ. Lett. 117 (1), 208–210.
Cameron, D.R., 1978. The expansion of the public economy: a comparative analysis. Am. Polit. Sci. Rev. 72, 1243–1261.
Canay, I.A., 2011. A simple approach to quantile regression for panel data. Econ. J. 368–386.
Ciccone, A., Jarocinski, M., 2010. Determinants of economic growth: will data tell? Amer. Econ. J. Macroecon. 2 (4), 222–246.
Dreher, A., Sturm, J.-E., Ursprung, H.W., 2008. The impact of globalization on the composition of government expenditures: evidence from panel data. Public Choice 134 (3–4), 263–292.
Durevall, D., Henrekson, M., 2011. The futile quest for a grand explanation of long-run government expenditure. J. Public Econ. 95 (7), 708–722.
Feenstra, R.C., Heston, A., Timmer, M.P., Deng, H., 2009. Estimating real production and expenditures across nations: a proposal for improving the Penn World Tables. Rev. Econ. Stat. 1, 201–212.
Feenstra, R.C., Inklaar, R., Timmer, M., 2011. The Next Generation of the Penn World Table. University of California, Davis and University of Groningen, (in process).
Feldkircher, M., Zeugner, S., 2012. The impact of data revisions on the robustness of growth determinants—a note on 'determinants of economic growth: will data tell?'. J. Appl. Econ. 27, 686–694.
Heston, A., Summers, R., Aten, B., 2002. Penn World Table Version 6.1. Center for International Comparisons at the University of Pennsylvania (CICUP), Baltimore, MD.
Johnson, S., Larson, W., Papageorgiou, C., Subramanian, A., 2013. Is newer better? Penn World Table revisions and their impact on growth estimates. J. Monet. Econ. 60 (2), 255–274.
Liberati, P., 2007. Trade openness, capital openness and government size. J. Public Policy 2, 215.
Neary, J.P., 2004. Rationalizing the Penn World Table: true multilateral indices for international comparisons of real income. Am. Econ. Rev. 1411–1428.
Ram, R., 2009. Openness, country size, and government size: additional evidence from a large cross-country panel. J. Public Econ. 93, 213–218.
Rodrik, D., 1998. Why do more open economies have bigger governments? J. Polit. Econ. 997–1032.
White, H., 1980. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. Econometrica 48, 817–838.