ECONOMIC ANNALS

EKONOMSKI ANALI, FOUNDED IN 1955
BY THE FACULTY OF ECONOMICS, UNIVERSITY OF BELGRADE
VOLUME LXVII, NO. 232 / JANUARY – MARCH 2022

232

UDC: 3.33 ISSN: 0013-3264

ECONOMIC ANNALS

Publisher: Faculty of Economics, University of Belgrade

For Publisher the Dean
Žaklina Stojanović

Editor-in-Chief
William Bartlett

Secretary of the Editorial Board
Ivana Ivković

Technical Assistance
Marina Milošević

Language Editor
Rowena Rowling

Cover Design
Milan Novičić

Editorial office and administration
FACULTY OF ECONOMICS
11000 Belgrade, Kamenička 6, Serbia
Tel: (381)(11) 3021-210
Fax: (381)(11) 2639-560

Website: *http://www.ekof.bg.ac.rs/publikacije/casopisi/ekonomski-anali/*
E-mail: *ea@.ekof.bg.ac.rs*

The journal is published quarterly
Annual subscription: 2400 RSD
Account No. 840-1109666-73
(Faculty of Economics, Belgrade)
Circulation: 200 copies

UDC: 3.33 • ISSN: 0013-3264

Print
JAVNO PREDUZEĆE „SLUŽBENI GLASNIK" – Beograd, www.slglasnik.com

ECONOMIC ANNALS 232 / 2022

ECONOMIC ANNALS, Volume LXVII, No. 232 / January – March 2022
UDC: 3.33 ISSN: 0013-3264

https://doi.org/10.2298/EKA2232007L

*Lara Lebedinski**
*Marko Vladisavljević*****

PARENTHOOD AND LABOUR MARKET OUTCOMES IN SERBIA

ABSTRACT: *Using the Labour Force Survey data for the period 2014 to 2018 for Serbia, this paper explores the effect of parenthood on the labour market trajectories of parents, the so-called 'parenthood penalty'. We find that mothers are less likely than non-mothers to be active in the labour market when their children are very young, but this effect is transitory, and mothers of older children are actually more likely to be active than non-mothers. Similarly, we observe that mothers of small children are less like-ly to work overtime than non-mothers, but also that both parents of older children are more likely to engage in overtime work than men and women without children. We find a motherhood penalty in terms of hourly wages for mothers with younger children, but the penalty is not significant as children become older. By contrast, fathers are more likely to be active than non-fathers. We do not find an effect of fatherhood on hours worked or hourly wages. Overall, our results suggest that the motherhood penalty is present in Serbia in the early stages when children are young, but motherhood does not seem to have lasting effects on the labour market participation, hours worked, or wage rates of mothers. We do not find evidence of a fatherhood bonus, but we find that fathers are more likely to be active than non-fathers.*

KEY WORDS: *motherhood penalty, female employment, household economics*

JEL CLASSIFICATION: J16, J13, J31.

* Institute of economic sciences, Belgrade, Serbia, email: lara.lebedinski@ien.bg.ac.rs
** Institute of economic sciences, Belgrade, Serbia; Faculty of Economics, University of Belgrade, Serbia email: marko.vladisavljevic@ien.bg.ac.rs

1. INTRODUCTION

Women in Serbia are in a worse position in the labour market than men: they are less likely to hold a job and their salaries are lower than those of their male counterparts. According to the Labour Force Survey, in 2019 the female labour market participation rate in Serbia was 47.1%, while the male participation rate was 62.8%. In 2015 the raw (unadjusted) gender hourly wage gap in Serbia was 5.7% and the adjusted gender hourly wage gap was 12.5%, both in favour of men (Anić and Krstić, 2019).

This paper aims to understand whether and to what extent childbirth affects the labour market position of parents in Serbia and whether it can be related to the gender inequality in the labour market. The evidence from other countries suggests that the motherhood penalty is responsible for at least some of the gender pay gap (Kleven et al., 2019). There are different margins where the difference between females and males can emerge. The extensive margin refers to the level of female labour force participation, the intensive margin refers to the number of hours worked, and there is also the difference in hourly wages. To understand how motherhood changes the labour market opportunities of women we first look at differences between mothers and non-mothers. We then separate mothers into three groups based on the age of the youngest child and compare their labour market outcomes with those of females without children. We rely on cross-sectional data from the Labour Force Survey for the years 2014 to 2018 and control for a large number of individual and regional characteristics. To compare how labour market outcomes change after childbirth for women and men we provide additional evidence on fathers, using the same specification as for mothers.

Our findings suggest that, on average, the labour force participation of mothers does not differ from the participation of non-mothers. However, there is heterogeneity among mothers based on the age of the youngest child. Mothers with younger children are less likely to be active in the labour market and the participation rate increases as their children become older. Mothers of children aged 7 to 15 years are actually more likely to participate in the labour force than non-mothers. In contrast to mothers, fathers are more likely to be active in the labour market than non-fathers and their participation rate does not depend on the age of the child. This suggests that fathers' engagement with children does not vary substantially with the age of the child. Regarding hours worked, we observe that

mothers of younger children work less than non-mothers, but as children become older mothers increase their hours and when their child is aged 7 to 15 they work even more than non-mothers. For males we find that fathers work significantly more hours than non-fathers when their child is in the age group 7 to15. We find a significant hourly wage penalty for mothers of very young children compared to non-mothers, but this difference becomes smaller and insignificant as the children grow older. No hourly wage penalty is observed for fathers.

We contribute to the literature on the motherhood penalty by studying the case of Serbia. Countries have different histories, traditions, and institutional settings, and it is important to understand how these different factors interplay and affect women's position in the labour market in specific settings. This study is the first in the literature to focus on whether motherhood is correlated with female labour market outcomes in Serbia. Our study of the case of Serbia shows that the motherhood penalty need not be persistent over time, in contrast to the literature focusing on Western Europe (Kleven et al., 2019). Additionally, and again contrary to the literature on Western Europe, we show that in Serbia non-mothers and non-fathers have lower participation rates than females and males with children, so lower caring responsibilities seem to enable more leisure time.

Our findings suggest two relevant policies for the Serbian context which should be further explored. First, increasing the availability of childcare for children below the age of three would help mothers of young children increase their labour force participation. Second, a share of parental leave exclusively reserved for fathers should be considered. It has been shown in other countries that this policy can increase fathers' involvement in childcare and housework and make the within-household division more equitable, which in turn can help women focus on paid work in the labour market.

This paper proceeds as follows. Section 2 reviews the related literature, section 3 summarizes the Serbian context, section 4 describes the dataset and provides the descriptive statistics, section 5 gives the methodology, section 6 describes our findings, and section 7 discusses the results and concludes.

2. LITERATURE REVIEW

Childbirth affects the labour supply of women in the short term and reduces their life-time earnings. A large number of papers study this phenomenon, known as the motherhood penalty. Differences in the labour market participation and wages of mothers and non-mothers are found in all OECD countries, but the extent varies (OECD, 2012). The two most important factors affecting the motherhood penalty are policies that affect the work–family balance and cultural norms.

Maternity and parental leave policies and availability of childcare are the most relevant policies shaping female labour supply. A lack of job protection after childbirth can push women out of the labour force (Blau and Kahn, 2013) if they have to decide between career and family. However, the duration of maternity leave should not be too long, because a long leave can worsen mothers' position in the labour market (Schönberg and Ludsteck, 2014). The availability of childcare is a necessary condition for women to be able to return to work. Both the availability and the price of childcare affect female labour supply. Childcare subsidies are an important and effective policy to incentivize women to return to the labour market after childbirth (Givord and Marbot, 2015; Simonsen, 2010), and the quality of available childcare also affects women's decision to return to work. If mothers know that formal childcare meets the needs of their child they will be more willing to use it and return to the labour market. Other policies such as paternity leave and lower marginal tax rates on second earners can also encourage women to return to the labour market (Budig et al., 2016).

Cultural norms shape the institutional setting in each country, but institutional factors cannot explain all the differences between countries in maternal labour force participation. Aside from the indirect effect of institutions, traditions and social norms directly affect mothers' decision of whether, when, and to what extent to return to the labour market. Mothers' employment, and hence the gender gap, is affected by the cultural roles of males and females in the household and at work. In fact, Budig et al. (2012) provide evidence from a cross-sectional study that cultural attitudes amplify associations between parental leave, publicly funded childcare, and maternal employment. If in a setting there is cultural acceptance of working mothers, then supportive policies reinforce maternal employment. However, policies favouring maternal employment are less effective in

conservative settings where mothers are expected to be responsible for childcare and housework.

Most papers studying the effect of parenthood on the gender pay gap discuss this phenomenon from the perspective of mothers. However, there is also evidence that fathers experience a "baby bonus" after childbirth (Hodges and Budig, 2010). There are three potential explanations in the literature for fathers having higher earnings than non-fathers. First, Becker's specialization hypothesis (Becker, 1981) posits that men specialise in market work while women specialise in household work. Alternatively, Gray (1997) provides evidence that more productive men sort into marriage. Evidence from European countries suggests that fathers spending more paternal time report higher earnings than fathers spending less time with their children. This suggests that fathers are either involved in both paid work and childcare or are not involved in either (Smith Koslowski, 2011). Lastly, fathers have higher expenses than non-fathers and have to earn more when children are born. As a result, fathers are incentivised to work and consequently earn more after childbirth.

In recent years, long panel datasets on earnings have become available to researchers, making it possible to study the income of mothers and fathers over a relatively long period after childbirth (e.g., 10 to 20 years). Bertrand et al. (2010) were the first to show how gender differences in earnings emerge after completion of education and at the onset of young professionals' careers. While males and females have similar earnings shortly after completing education, they find that 10 to 16 years after completing an MBA, males have an advantage of 60 log points in terms of earnings. Kleven et al. (2019) use long-spanning panel data to estimate long-term cumulative earnings after childbirth for both fathers and mothers. The authors aim to understand to what extent motherhood and the motherhood penalty (or the equivalent child penalty for mothers) can explain the persistent gender inequality in the labour market. They examine three margins in which mothers can experience penalties in earnings: the extensive margin of labour supply (employment), the intensive margin of labour supply (hours worked), and the wage rate. They find that in the Scandinavian and Germanic countries the extensive margin effects are smaller than the earnings effects, while in the Anglo-Saxon countries (the UK and US) the employment penalty is the main driver of the earnings penalty. Kleven et al. (2020) use the methodology from Kleven et al. (2019)

to study 60 years of parental leave and childcare policy experimentation in Austria. Surprisingly, they find that parental leave and childcare policies do not reduce the gender gap. They argue that gender inequalities are driven by equilibrium features of the labour market and not by public policies.

Cross-country evidence on the motherhood and parenthood penalty is relatively scarce, and estimates are typically based on the data for one country. One important recent study is a meta-analysis of the motherhood penalty by Cukrowska-Torzewska and Matysiak (2020). They analyse studies estimating the motherhood penalty and find that the average motherhood wage gap is around 3.6% to 3.8%. They also find that the residual gap in wages is smallest in Nordic countries, slightly larger in Belgium and France, and largest in "post-socialist countries of Eastern and Central Europe" (specifically Poland and Ukraine) and Anglo-Saxon countries. They also stress that women in the post-socialist countries have among the lowest employment rates in Europe, but despite low childcare availability return to full-time work relatively quickly.

Other evidence on post-socialist countries suggests that after childbirth women in Russia initially experience strong employment penalties (a reduction in employment levels of between 40 and 65 percentage points), and while these penalties are lasting they stabilise at about 6% after five years (Lebedinski et al., 2020). The same study finds no penalties in terms of working hours or hourly wages. The authors explain these findings in terms of the limited availability of non-standard employment options such as part-time jobs, which are a mechanism frequently used in the EU and US to balance family and work life. The authors conclude that in Russia the options for women are limited to either completely withdrawing from the labour force or returning to their previous work.

3. CONTEXT

In Serbia the fertility rate has been falling since the 1990s. In 2020 the average age of first-time mothers was 28.7 years and in 2018 the fertility rate was 1.5 (SoRS, 2019a). Compared to other European countries, the length of maternal and parental leave and the monetary compensation in Serbia are generous. The maternity leave period starts 45 to 28 days before the due date and lasts 3 months. Maternity leave is followed by a parental leave period which lasts 9 months. Maternity leave can only be taken by mothers, while parental leave can be taken by one

of the two parents or shared between them. The compensation for working parents during maternity and parental leave is 100% of monthly average earnings in the 18 months preceding the leave. After the birth of a child, fathers get paid leave of up to 5 working days. While fathers can take parental leave, it is rare that they do so. In 2019, out of a total of 64,399 births, only 328 fathers took parental leave.

One important factor determining the participation rate of mothers of young children is the availability and quality of childcare. Children in Serbia can enter childcare at 6 months. In 2019 the enrolment rate was 28.1% for children aged 0 to 2 years, and 66.4% for children aged 3 to 5 years (excluding the compulsory preschool programme) (SoRS, 2020b). The enrolment rate in compulsory preschool education from age 6 to 7 was 97.4% in 2019 (SoRS, 2020b). The childcare enrolment rate of children aged 0 to 2 is similar to the OECD average (35% in 2017: OECD, 2020), but the preschool enrolment rate is more than 20 percentage points below the OECD average (87.2% in 2017). Public childcare facilities and preschools in Serbia are oversubscribed and availability of preschool places can be an obstacle to female labour force participation. In larger cities this problem could be partly solved by the provision of private preschool education vouchers to families that cannot get places in public preschools (SoRS, 2021). However, there is still excess demand for kindergartens, which makes it more difficult for women to search for and take a job. Preschool education quality is frequently measured by the child-to-teaching-staff ratio.[1] In 2018 the average child-to-teaching-staff ratio for the 3 to 5 year age group was 14.2 in OECD countries and 11.6 in Serbia (OECD, 2020; SoRS, 2019b). Using the child-to-teaching-staff ratio as a proxy for quality, the Serbian preschool education is somewhat better than OECD average.

Finally, let us briefly discuss the social and gender norms in Serbia. Although Serbia is a former communist country with a high female labour force participation rate, it is also a country where most of the housework and child-rearing traditionally falls to the females in the household. Data from 2015 suggest that women in

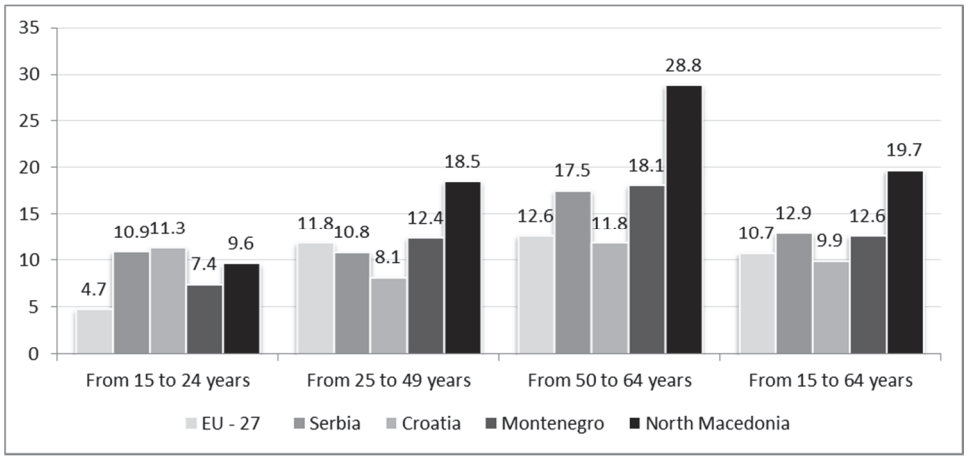
[1] This indicator does not take into account that teachers are not present throughout the working day. For instance, in Serbia, each full-time educational group has 2 full-time teachers and each of them spends 6 hours with the group and they both usually cover the period from 8am to 5pm. The two teachers actually overlap for only 3 hours.

Serbia spend 4 hours and 48 minutes on the household and family per day, while men spend 2 hours 33 minutes (Eurostat, 2019).

In the 2016 Gender Equality Index, Serbia was in the lowest tercile in Europe with a score 10.4 points lower than the EU average (Babović, 2018). The largest differences are in the domain of money, where Serbia's score is 19.4 points lower than EU average, mainly due to elderly women, women living in rural areas, women living in single households, and single mothers all being at high risk of poverty; and in the domain of time, where Serbia's score is 17 points lower than the EU average due to a significantly higher burden of housework and less time available for recreation and participation in cultural or social activities. The lowest difference with the EU is in the labour market, the focus of this paper, where Serbia is only 3.3 points below the EU average, mainly due to small participation gaps and relatively favourable work quality (Babović 2018). For example, women are employed on permanent contracts more frequently than men.

Figure 1 presents the gender employment gap in the EU, Serbia, and selected neighbouring countries in 2019. The total employment gap for the working-age population (15–64) in Serbia of 12.9 percentage points (pp) is 2 pp higher than the EU average of 10.7 pp. However, the employment gap for the 25 to 49 age group, which is closest to the sample that we use to estimate the motherhood penalty (25–45 years), is slightly lower in Serbia (10.8 pp) than the EU-27 average (11.8 pp). Therefore, on average, women aged 25 to 49 in Serbia are not in a less favourable position than women in the EU-27. A similar trend can be observed in all the countries in the region: compared to the EU the difference in gender employment gap is lower for the 25 to 49 age group, and higher for younger and older workers.

Figure 1: Gender employment gap in EU, Serbia, and selected countries, 2019, in percentage points

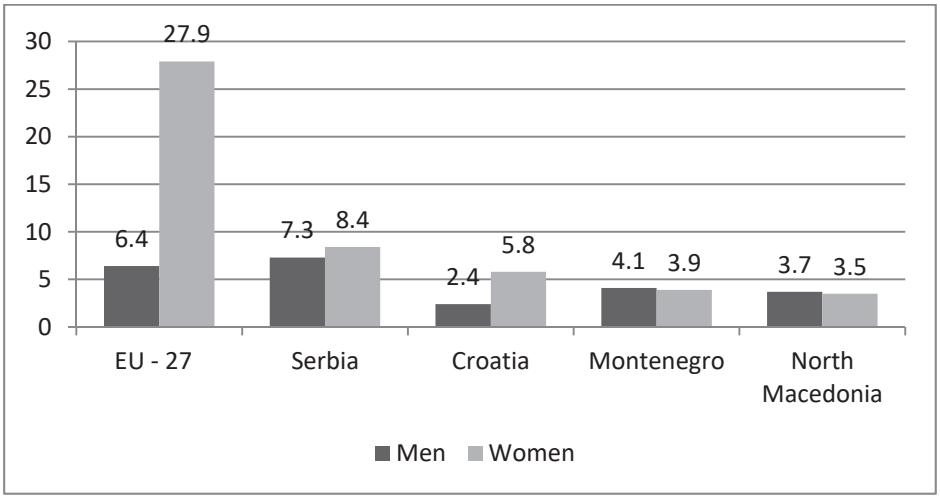


Source: Eurostat database: Labour Force Survey (lfsa_ergan indicator)

Previous research indicates that the lower employment of women in Serbia compared to men is associated with low employment opportunities among women with low levels of education, the higher disincentivizing impact of receiving social transfers, and the presence of young children in the household (Žarković-Rakić & Vladisavljević, 2016)

Figure 2 indicates that women work part-time in Serbia much less than in the EU. Part-time work is frequently used in European countries to achieve family–work balance and to re-integrate mothers in the labour market. In the EU, 27.9% of employed women age 25–49 work part-time, while in Serbia this share is only 8.4%. On the other hand, the share of male part-time workers aged 25–49 in Serbia is 7.3%, slightly higher than in the EU-27. Low shares of part-time work are also characteristic of other countries in the region.

Figure 2: Share of part-time employment by gender in EU, Serbia, and selected countries, 2019, %, age group 25–54



Source: Eurostat database: Labour Force Survey (lfsa_eppgan indicator)

A more disaggregated analysis shows that part-time work in Serbia is predominantly in the agriculture, forestry, and fishing sector, and in households as employers and undifferentiated goods and service-producing activities of households for their own use, where almost two-thirds of workers are employed part-time. In the EU the distribution of part-time work across sectors is much more uniform, with the two sectors dominant in Serbia presenting only 8% of total part-time employment. Previous research indicates that in Serbia part-time work for both genders is much more frequent in informal than formal employment (Žarković-Rakić & Vladisavljević, 2016).

In 2015 the unadjusted gender gap in hourly wages in Serbia was 5.7% (Anić & Krstić, 2019). The gap is relatively stable: it was 6.2% in 2008 and 3% in 2011 (Avlijaš et al., 2013). However, while on average women's wages are lower they have better labour market characteristics than men: higher levels of education, a more favourable occupational structure, and higher share of public sector employment (Vladisavljević et al., 2015). When controlling for these characteristics, the estimated adjusted wage gap in Serbia is higher than the simple difference in

wages. Different estimates from 2008 to 2015 suggest that the adjusted gap is between 9% and 15% (Avlijaš et al., 2013; Žarković-Rakić & Vladisavljević, 2016; Anić & Krstić, 2019).

All the above studies primarily focus on the estimation and explanation of the gender gap in wages and do not explicitly analyse the impact that children have on the wages of mothers and fathers. They account for the impact of children implicitly, typically by using this variable in the selection equation and correcting the wage equation for the effects of the selection.

4. SAMPLE AND VARIABLE DESCRIPTION AND DESCRIPTIVE STATISTICS

This study uses the Labour Force Survey (LFS) for Serbia for the years 2014 through 2018. The LFS is a nationally and regionally (NUTS2 level) representative continuous[2] survey and its goal is to monitor the labour market situation and to deliver internationally established and comparable indicators, such as employment and unemployment rates. The LFS sample is a two-stage stratified sample, with the 2011 Serbian Population Census frame used as a sample frame for the selection of enumeration areas, as first-stage sampling units and households and as second-stage sampling units (SORS, 2020a). The LFS is conducted by the Statistical Office of the Republic of Serbia.

The LFS does not ask explicitly whether adults have children and if so how many. However, the LFS household roster does collect basic socio-demographic information on all household members, including questions on the identification numbers of the child's mother and father (or legal guardian), which enables us to link information on children with their parents if they live in the same household. The LFS is conducted both in-person and over the phone; however, the household roster and basic socio-demographic information on all household members is collected in person by interviewers (SoRS, 2017), so we can be confident that our methodology identifies households with children and that household rosters are a reliable source of information on household members.[3] Using this information,

[2] The survey has been continuous since 2015 in Serbia.

[3] According to LFS estimates from 2018 there were 993,843 children aged 0 to 14 in Serbia. This is very similar to the data provided in the Demographic Yearbook, which suggests that in the same year there were 1,000,596 children (SoRS, 2019a).

we define three groups of interest: 1) mothers and fathers who have own children in the household and whose youngest child is 15 or younger, 2) non-mothers and non-fathers who do not have children (of any age), and 3) parents whose youngest child is older than 15 years, who we exclude from the analysis. The latter decision is based on the fact that older children are more able to take care of themselves and to assist in household chores. Additionally, older children are more likely to have moved out of the household, which increases the likelihood of classifying their parents as non-parents, where instead they should be dropped from the sample as parents of children older than 15.

For the purpose of this study, we restrict the sample to individuals aged 20 to 50 years. We set the lower age limit at 20 for two main reasons. First, there are very few births among women and men younger than 20 years old,[4] and second, individuals below this age are largely still in education (mostly high-school) and so are inactive in the labour market and overwhelmingly do not want to work, indicating that their labour supply is inelastic (Arandarenko et al., 2012). This is particularly true for Serbia, where small jobs that could be performed during education are practically non-existent and in general it is very difficult to balance education and work, so the determinants of their labour supply and wages would be different from those of the general population and bias the regression analysis results.

On the other hand, the decision to set the upper limit at 50 years is motivated by concerns about classifying persons in our sample as non-parents, instead of excluding them from our sample (as parents of children older than 15 years). We can only know that someone is a parent if they live in the same household as their child and we exclude from the analysis parents of children older than 15 years. Since on average there is a strong positive correlation between the age of the parents and the ages of their children (the older the parents the older their children), and since older children are more likely to have moved out of the parental home, it is reasonable to assume that the error of classifying persons as non-parents instead of parents of (absent) older children increases with age of the parents. The average home-leaving age in Serbia was between 30 and 31 years in 2014–2018

[4] The oldest child in our sample was born in 1999 and the youngest in 2018. In 1999 out of all births, only 9.2% of mothers were younger than 20 years (SoRS, 2006). In 2018, 3.8% of births were delivered by mothers younger than 20 years (SoRS, 2019a).

(Eurostat, 2021), largely due to poor financial situation (Milić and Zhou, 2015). This relatively late home-leaving age makes it possible to also include older women in our analysis, e.g., aged 45 to 50, because it is unlikely that their children have left the household.

Table 1 provides descriptive statistics for the sample of mothers and non-mothers jointly and separately. We can see that mothers are somewhat older and less-educated than non-mothers. Most mothers have two children (47.2%), followed by one child (41.1%), while a few have more than two children (11.7%). Non-mothers live with fewer adults in the household than mothers, but this difference is not large. Both mothers and non-mothers live in households with 2 to 3 household members, because many non-mothers still live with their parents. As expected, non-mothers are more prevalent in economically developed parts of the country, namely Belgrade and other urban areas.

Regarding labour market outcomes, mothers are more likely to be both active (74.2% of mothers and 62.9% of non-mothers) and employed (61.1% of mothers and 45.9% of non-mothers). That mothers are more likely to be employed than non-mothers is somewhat unexpected, but could be explained by the composition of the household and mothers' lower reservation wage. The simple comparison shown in Table 1 does not take into account the socio-demographic characteristics of mothers and non-mothers, and these characteristics could explain their different outcomes (differences in age, educational background, etc.) Non-mothers earn a marginally higher monthly salary than mothers, as they are more frequently among those with wages higher than 45,000 RSD (about 14.7% of non-mothers, as opposed to about 13.5% of mothers). At the same time, there is no difference between mothers and non-mothers in the average working hours per week. The vast majority of non-mothers live in households with at least one parent (68.0%) while this is the case for only a small fraction of mothers (10.1%). As a result, mothers have higher expenses and are willing to accept a lower wage, as confirmed by the reservation wage.

Table 1: Descriptive statistics: Mothers and non-mothers

	Total N=29,939	Mother N=13,953	Non-mother N=15,986	p-value
Socio-demographic characteristic				
Age	32.70 (±8.41)	34.80 (±6.32)	30.86 (±9.50)	<0.001
Highest completed educational level				<0.001
Primary school or less	13.1%	16.0%	10.5%	
General or VET secondary school	58.8%	57.3%	60.1%	
College, university, or higher	28.2%	26.8%	29.4%	
Married	48.3%	81.8%	19.1%	<0.001
Number of children				<0.001
No children	53.4%	0.0%	100.0%	
1 child	19.2%	41.1%	0.0%	
2 children	22.0%	47.2%	0.0%	
3 or more children	5.4%	11.7%	0.0%	
Number of adults in household	2.48 (±1.08)	2.67 (±1.09)	2.32 (±1.05)	<0.001
Nuts 2 level				<0.001
Belgrade	25.1%	23.1%	26.9%	
Vojvodina	25.0%	26.0%	24.0%	
Šumadija and Western Serbia	27.2%	27.5%	26.9%	
Eastern and Southern Serbia	22.8%	23.4%	22.2%	
Urban	61.1%	59.1%	62.9%	<0.001
Labour market outcomes				
Active	68.2%	74.2%	62.9%	<0.001
Employed (SoRS)	53.0%	61.1%	45.9%	<0.001
Monthly net wage				<0.001
Less than 17,000 RSD	5.0%	4.2%	6.0%	
More than 17,001 and less than 25,000 RSD	32.3%	33.0%	31.5%	
More than 25,001 and less than 35,000 RSD	29.8%	29.6%	30.1%	
More than 35,001 and less than 45,000 RSD	18.9%	19.8%	17.8%	
More than 45,001 and less than 60000 RSD	9.7%	9.4%	10.1%	
More than 60,001 and less than 80,000 RSD	2.6%	2.6%	2.7%	
More than 80,001 RSD	1.7%	1.5%	1.9%	
Usual numbers of hours worked in a week	41.66 (±6.52)	41.76 (±5.96)	41.55 (±7.09)	0.060

Notes: Data are presented as mean (±SD) for continuous measures, and % for categorical measures.

Table 2 compares fathers and non-fathers. Similar to mothers, fathers are somewhat older and more educated than non-fathers. Most fathers have two children (49.9%), followed by one child (37.8%), while the rest have three or more children (12.3%). In terms of regional distribution there are only small differences between fathers and non-fathers. There are no differences between fathers and non-fathers living in rural and urban settings. Fathers are both more likely to be active and more likely to be employed than non-fathers. Fathers earn more than non-fathers, but they do not work more hours. The lower activity and likelihood of being employed among non-fathers can be explained by household composition and the lower level of expenses: 81.5% of non-fathers live with their parents, whereas this is the case for only 40.3% of fathers.

The labour market outcomes of women depend on the ages of the children.[5] Mothers with younger children generally have more difficulty reconciling work and childcare and therefore they are less likely to be part of the labour force.[6] As children get older, mothers return to work, and this is also confirmed in the Serbian case. Figure 3 shows the participation rate of females and males based on the age of the youngest child. We observe that women with children aged 0 to 2 years have a similar participation rate to non-mothers and the lowest participation rate among mothers. As children age, the likelihood of entering the labour market for women increases. Notably, women with children aged 7 to 15 years have the highest participation rate among all four groups. In contrast to women, men with children have a considerably higher labour market participation rate and the age of the child does not correlate with the likelihood of men being active in the labour market.

5 See, for instance, Grimshaw and Rubery (2015).

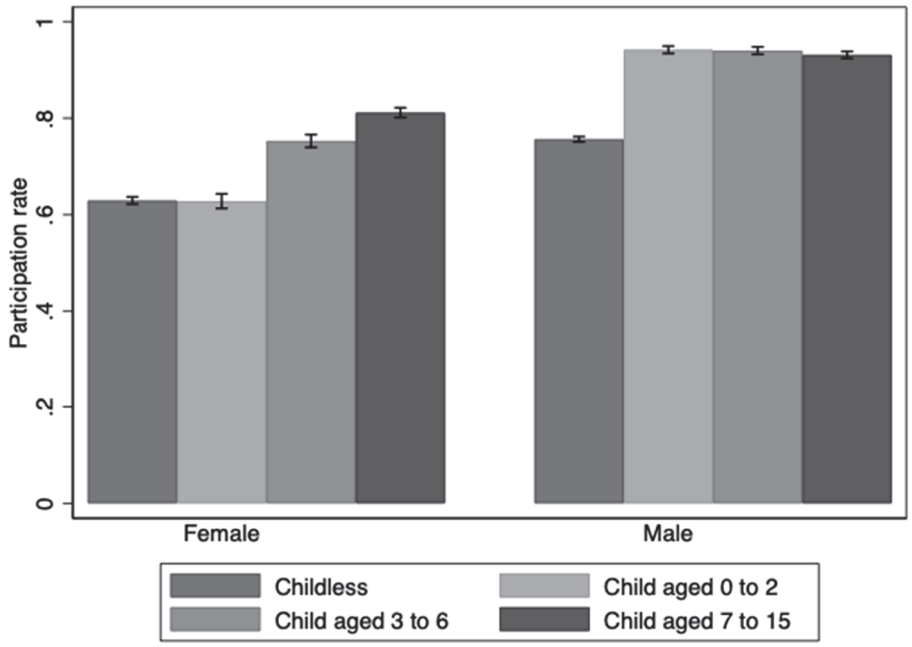
6 Note that the labour force consists of all individuals who are either employed or actively searching for work; i.e., unemployed as defined by the International Labour Organization. The labour force is considered to be the active population.

Table 2: Descriptive statistics: Fathers and non-fathers

	Total N=35,581	Father N=12,168	Non-father N=23,413	p-value
Socio-demographic characteristic				
Age	33.44 (±8.48)	37.62 (±6.05)	31.26 (±8.74)	<0.001
Highest completed educational level				<0.001
Primary school or less	14.7%	15.8%	14.1%	
General or VET secondary school	67.9%	65.4%	69.2%	
College or university or higher	17.4%	18.8%	16.7%	
Married	36.4%	88.2%	9.5%	<0.001
Number of children				<0.001
No children	65.8%	0.0%	100.0%	
1 child	12.9%	37.8%	0.0%	
2 children	17.1%	49.9%	0.0%	
3 or more children	4.2%	12.3%	0.0%	
Number of adults in household	2.45 (±1.09)	2.75 (±1.07)	2.30 (±1.07)	<0.001
Nuts 2 level				0.079
Belgrade	22.9%	22.2%	23.3%	
Vojvodina	25.0%	25.5%	24.7%	
Šumadija and Western Serbia	28.1%	28.1%	28.2%	
Eastern and Southern Serbia	24.0%	24.3%	23.8%	
Urban	57.2%	57.4%	57.0%	0.49
Labour market outcome				
Active	81.8%	93.7%	75.6%	<0.001
Employed (SoRS)	66.1%	83.0%	57.4%	<0.001
Monthly net wage				<0.001
Less than 17,000 RSD	4.5%	3.1%	5.6%	
More than 17,001 and less than 25,000 RSD	23.7%	20.6%	26.1%	
More than 25,001 and less than 35,000 RSD	34.0%	31.9%	35.5%	
More than 35,001 and less than 45,000 RSD	19.1%	21.1%	17.6%	
More than 45,001 and less than 60,000 RSD	12.4%	15.4%	10.0%	
More than 60,001 and less than 80,000 RSD	3.6%	4.7%	2.8%	
More than 80,001 RSD	2.7%	3.3%	2.3%	
Usual numbers of hours worked in a week	43.26 (±8.03)	43.34 (±7.60)	43.20 (±8.33)	0.27

Notes: Data are presented as mean (±SD) for continuous measures, and % for categorical measures.

Figure 3: Labour force participation, females and males, by age of youngest child

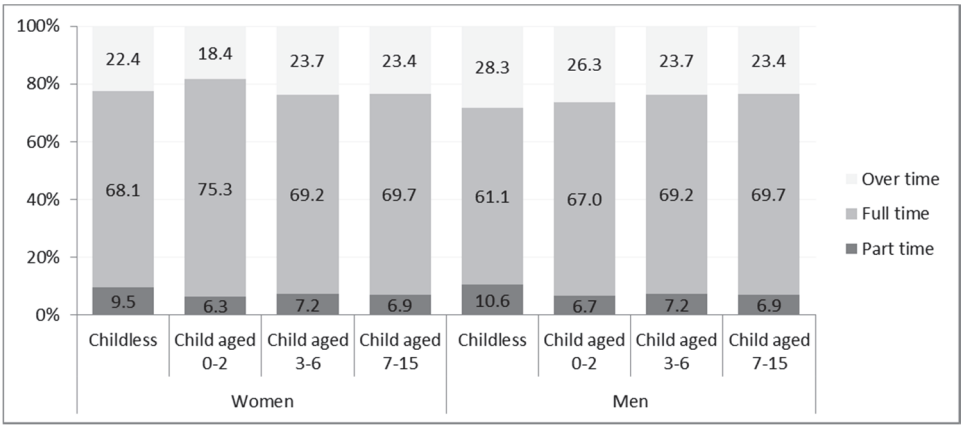


As discussed previously, part-time work in Serbia is rare, with the share of part-time workers in the sample being about 8%. The detailed distribution of the usual hours worked per week (Figure A1 in the Appendix) suggests that the distribution of working hours is highly discrete, with two peaks. More than 60% of women (both mothers and non-mothers) work 40 hours per week, while another 20% of women work 48 hours per week. Similarly, between 50% and 60% of both fathers and non-fathers work 40 hours per week, and approximately another 25% work 48 hours. Based on this distribution we divide all workers into three groups: 1) part-time workers (those working less than 35 hours per week), 2) full-time workers (those working between 35 and 44 hours per week), and 3) overtime workers (those working more than 45 hours per week).

Figure 4 presents the share of workers working part-time, full-time, and overtime, by gender and age of their youngest child. On average, about 8% of the sample works part-time, about two-thirds work full-time, and about a quarter work overtime. Overall, the differences between the groups are not prominent. Contrary to

expectations, both childless men and childless women have slightly higher shares of part-time workers than all mothers and fathers. On the other hand, women with small children work overtime less frequently (about 18% of cases) than other groups (on average about 23%). This is expected, as due to increased responsibilities at home taking care of the infant they cannot work additional hours when they return to work. Furthermore, men with children aged 3 to 6 and 7 to 15 work overtime slightly less frequently than childless men and men with small children.

Figure 4: Hours worked per week, females and males, by age of youngest child



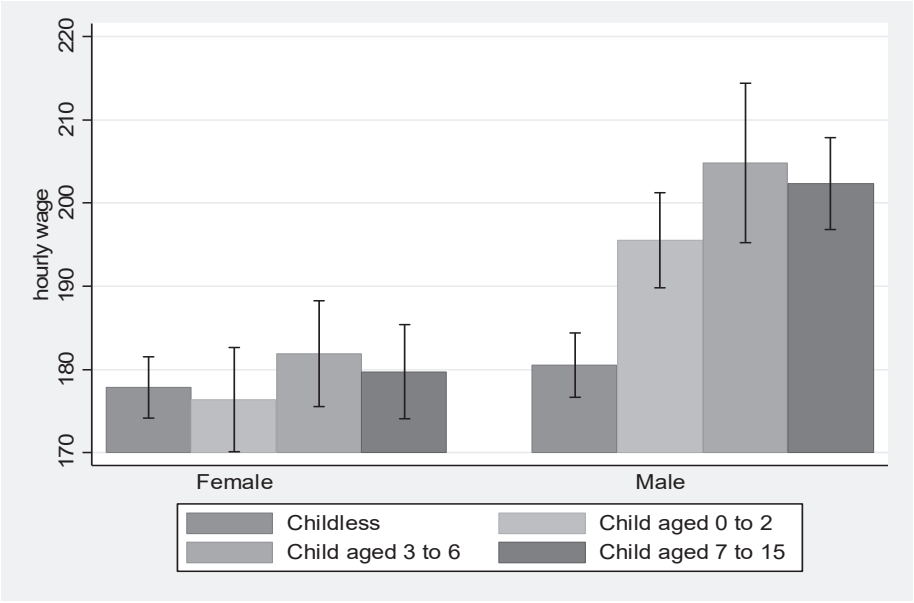
Data on wages in LFS are collected at the net monthly level, only for wage-employed workers, i.e., employees.[7] The wages are then transformed into hourly wages[8] and inflated to 2018 levels using the Consumer Price Index (2018=100). Figure 5 indicates that on average men have higher wages than women. There are no significant differences between mothers and non-mothers. On the other hand,

[7] Employees are first asked to provide the exact amount of monthly wages earned in the previous month. If they are not able or willing to provide an answer, they are asked to provide an interval for the monthly wages earned. In order to compute hourly wages, the exact amount of monthly wages is required. The potential bias in the wage equation that could occur due to omitting interval wages is accounted for by selection correction introduced in the wage equation. In total, about 61.4% of all workers who provided information on wages provided exact wages (including both those providing exact and interval wages).

[8] We divide the amount of monthly wages by 23/5 (average number of working days in a month/week) to arrive at weekly wages, and weekly wages with usual weekly working hours

fathers have higher wages than non-fathers, and the difference is most pronounced for fathers of children aged 3–6 years. In the next chapter we explain the methodology used to investigate if these hourly wage differences remain statistically significant when controlling for other relevant characteristics such as education, age, and region.

Figure 5: Hourly wage, females and males, by age of youngest child



5. METHODOLOGY

In the previous section we reported descriptive statistics of characteristics and outcomes for mothers and non-mothers and for fathers and non-fathers. The background characteristics of the groups differ both within the female group and within the male group, and in order to control for these differences in background characteristics we turn to multivariate regression analysis. This type of analysis enables us to condition on observable differences and to estimate the different labour market outcomes for mothers and non-mothers, and fathers and non-fathers.

Our baseline model comparing labour market outcomes of mothers and non-mothers relies on the following regression:

$$Y_{ri} = \alpha_0 + \alpha_1 \text{mother} + \boldsymbol{X_{ri}}\boldsymbol{B} + \mu_r + year + u_{ri} \quad (1)$$

where Y_{ri} is the outcome of interest and we consider three outcomes: participation in labour market, and, for those employed, hours worked per week and hourly log salary. Our coefficient of interest in regression (1) is the difference in outcomes between mothers and non-mothers captured by the coefficient α_1. The vector X_{ri} controls for background characteristics of the mother such as age, age squared, educational level, number of adult members in household, number of children, and whether the person lives in an urban or rural area. Additionally, for hours and wage equations we control for differences in the following job characteristics: occupation,[9] sector,[10] type of ownership (public or private), supervising position, firm size, and type of contract (permanent or temporary)[11]. Finally, μ_r

[9] We use ISCO 1-digit categorization of occupations. Category 10 – Armed forces occupations – is combined with category 2 – Professionals – due to small sample size.

[10] Based on NACE Rev. 2 classification of sectors, we group the employees into three groups working in: agriculture (sector A), industry (sectors B to F), and services (sectors G to U).

[11] In the case of the hours and wage equation, there is a potential bias in the estimates caused by non-random sample selection (Heckman, 1979). According to Heckman, sample selection bias can be viewed as the omitted variables problem, and resolved by adding a variable that represents the different characteristics of persons in the sample and persons not in the sample. Since the variables in the participation equation are also in the hours and wage equation, the exclusion restriction condition cannot be fully satisfied (at least one variable has to appear in the participation equation that is not in the hours equation). In this case it is more reasonable to adopt a model without correction, as suggested by Puhani (2000). However, initial estimates of the wage and hours equations suggested that some of the variables are not significant (see tables A2 and A3), while they are significant in the participation equation. We therefore drop the insignificant variables from the wage and hours equations and leave them in the participation equation, which enables us to fulfill the exclusion restriction condition. In this approach we first estimate the participation equation conditional on age, age squared, level of education, marital status, and number of adults and children in the household, via probit estimate. We then calculate the Inverse Mills Ratio (IMR), as a ratio of the probability density function and the cumulative probability distribution function, where the respective probability functions are derived from the participation equation (Wooldridge, 2002). This variable, according to Heckman, represents the differences in unobserved characteristics between wage employed and other groups in the labour market. Finally, we add IMR to the list of covariates in Equations (1) and (2). However, as the insignificant variables are still theoretically associated with hours of work (particularly in our framework), we opt to present the results without selection as our main results, and use estimates with selection as a robustness check.

are region fixed effects, *year* are survey year fixed effects, and u_{ri} is the random error term.

In a similar spirit, we estimate the labour market outcomes for non-mothers and the three different categories of mothers based on their youngest child:

$$Y_{ri} = \beta_0 + \beta_1 \text{mother}_{0-2} + \beta_2 \text{mother}_{3-6} + \beta_3 \text{mother}_{7-15} + \boldsymbol{X_{ri}}\boldsymbol{B} + \mu_r + year + u_{ri} \quad (2)$$

In the case of Equation (2) we are interested in coefficients β_1 , β_2, and β_3, which capture the differences between the non-mothers and mothers with children of different age groups. For instance, when estimating participation, β_1 captures whether mothers of children aged 0 to 2 years have a different participation rate than non-mothers. A positive coefficient β_1 would imply that mothers of children aged 0 to 2 are more likely to participate in the labour market, while a negative coefficient would mean that they are less likely to participate than non-mothers.

In the same way we estimate the labour market outcomes of fathers, first by grouping them together as we do for mothers in Equation (1), and then by separating them into three groups based on the age of the youngest child, as shown in Equation (2).

6. ESTIMATION RESULTS

6.1 Participation rate

As a first step towards understanding the situation of parents versus non-parents in the labour market, we estimate the participation rate and report it in Table 3. For mothers we find that there are no statistically significant differences between mothers and non-mothers, as shown in column (1) of the table. We then proceed to examining the heterogeneity among mothers based on the age of the youngest child, reported in column (2), and here it is clear that mothers of young children are less likely to be active in the labour market. Mothers with the youngest child aged 0 to 2 years are 6.4 percentage points less likely to be active than non-mothers. This is the expected result because within the household the mothers is responsible for child-rearing in the early years of a child's life and this is the period when they step back from the labour market. Mothers whose youngest child is 3

to 6 years old are not less likely to be active than non-mothers and, lastly, mothers whose youngest child is aged 7 to 15 are more likely to be active than non-mothers. One reason why mothers of older children are more likely to be in the labour market is that families with older children have higher expenses which cannot be covered by a sole earner. In all regressions we include individual characteristics affecting the propensity to be active in the labour market, regional fixed effects, and survey year fixed effects. The covariates have the expected signs and they are reported in the Appendix in Table A.1. Older and more-educated mothers are more likely to be active. The coefficient of age squared is negative and statistically significant, suggesting that prime-age mothers are the most active. The likelihood of being active in the labour market is the same for mothers with one and two children, but it is reduced for mothers with three or more children. Mothers in Belgrade and Šumadija and Western Serbia are more likely to be active than mothers from Vojvodina and Eastern and Southern Serbia.

Fathers are 8.6 percentage points more likely to be active than non-fathers. This is in line with the literature, which suggests that fathers are more likely to be working and that they earn more than non-fathers (Hodges and Budig, 2010). In terms of covariates (reported in Table A.1 in the Appendix) in the father regression, we find that older, more-educated, and married fathers are more likely to be active. The number of adults in the household is not statistically significant, but, interestingly, the propensity to be active falls with the number of children, and this is especially the case for fathers with three or more children. Fathers in rural areas and fathers residing in the poorest Eastern and Southern regions of Serbia are less likely to be active.

Table 3: Labour market participation of parents and non-parents

	Women		Men	
	(1)	(2)	(3)	(4)
Child 0 to 15	–0.002		0.086***	
	(0.008)		(0.009)	
Youngest child 0 to 2		–0.064***		0.105***
		(0.010)		(0.011)
Youngest child 3 to 6		0.006		0.083***
		(0.010)		(0.011)
Youngest child 7 to 15		0.041***		0.071***
		(0.009)		(0.010)
Individual-level controls	X	X	X	X
Regional FE	X	X	X	X
Survey year FE	X	X	X	X
Observations	29,669	29,669	35,311	35,311

Notes: Marginal effects from probit model. Individual-level controls: age, age squared, indicators for educational level, indicator for marriage status, number of adults in household, number of children in household, indicator for urban settings. Regional fixed effects: Belgrade, Vojvodina, Šumadija and Western Serbia, Eastern and Southern Serbia. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

6.2 Hours penalty

As discussed in Section 4, the distribution of working hours in Serbia is highly discrete, with over 80% of both men and women working 40 or 48 hours. Therefore, instead of analysing the continuous working hours variable, we analyse a categorical variable representing a choice between working part-time, full-time, and overtime. Due to the nature of the dependent variable, which represents three categories that can be ranked, we apply an ordinal probit estimator. Table A.2 in the Appendix presents the results of the hours worked equation. We first discuss the effects of the covariates and then turn our attention to the effect of parenthood. The results indicate that working hours are, ceteris paribus, higher for men and women with low education, in rural areas, and in regions other than Belgrade. Working hours increase for both genders linearly with age. Additionally, married women (after controlling for children) are working fewer working hours. Regarding job characteristics, workers in occupations such as sales and services and crafts have longer working hours, while longer working hours are

also observed for men working as clerks, plant and machine operators, in elementary occupations, and in industry or the services sector. Working hours are, ceteris paribus, shorter for the public sector, those working informally, and temporary workers. Additionally, men in supervisory positions work longer hours.[12, 13]

We now turn to the main focus of our paper – the effects of parenthood, which are presented in Table 4. The overall effect of motherhood is not significant, indicating that the working hours of mothers and non-mothers do not differ on average (column 1). However, when we split the motherhood effect by age of the youngest child, the results indicate that mothers of young children (youngest child aged 0 to 2) work less, while mothers of older children (7 to 14 years old) work longer hours than non-mothers (column 2). Mothers of children aged between 3 and 6 years are not significantly different from non-mothers in this respect. On average, the working hours of fathers and non-fathers do not differ (column 3), and no difference is found for fathers by age of the youngest child (column 4).

[12] Due to the fact that all the variables in the participation equation are also in the hours equation, the exclusion restriction condition (at least one variable has to appear in the participation equation that is not in the hours equation) for the application of the selection equation is not fulfilled. In this case it is more reasonable to adopt a model without correction, as suggested by Puhani (2000). However, the estimates in Table A.2 suggest that some variables that enter the model are not significant (age and age squared for women, age squared and number of children for men), while they are significant in the participation equation. Therefore, dropping the insignificant variables from the hours equation enables us to fulfill the exclusion restriction condition. Using this approach, we estimate the hours equation with the selection correction. The results from this model (available upon request) are very similar to the ones presented here and reaffirm our conclusions. However, as the insignificant variables are still theoretically associated with hours of work (particularly in our framework), we opt to present the results without selection as our main results.

[13] The results presented in Table A2 and Tables 4 and 5 refer to age group 25 to 45. An additional robustness check of these results was performed by including persons aged 20 to 50 years. Results, available upon request, yield similar coefficients and confirm the conclusions presented here.

Table 4: Parents' and non-parents' hours worked per week

	Women		Men	
	(1)	(2)	(3)	(4)
Youngest child 0 to 15	0.022		0.019	
	(0.032)		(0.033)	
Youngest child 0 to 2		–0.098**		–0.012
		(0.041)		(0.039)
Youngest child 3 to 6		0.018		0.032
		(0.040)		(0.039)
Youngest child 7 to 15		0.076**		0.039
		(0.035)		(0.038)
Individual-level controls	X	X	X	X
Job characteristics	X	X	X	X
Regional FE	X	X	X	X
Survey year FE	X	X	X	X
Observations	13,243	13,243	17,307	17,307

Notes: Results from ordinal probit model. Individual-level controls: age, age squared, indicators for educational level, indicator for marriage status, number of adults in household, number of children in household, indicator for urban setting. Job characteristic controls: occupation, sector, ownership, informal employment, type of contract, supervising position, and firm size. Regional fixed effects: Belgrade, Vojvodina, Šumadija and Western Serbia, Eastern and Southern Serbia. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

As the size coefficients in the ordinal probit model have no meaningful interpretation, we compute marginal effects for mothers of youngest children (aged 0 to 2) and mothers and fathers of older children (aged 7 to 15). These results are presented in Table 5. The marginal effects indicate that compared to non-mothers, mothers of youngest children (aged 0 to 2) are more likely to work part-time (by about 1 percentage point) or full-time (by 1.8 percentage points) and less likely to work overtime (by 2.7 percentage points). On the other hand, mothers of older children (aged 7 to 15) are more likely to work overtime than non-mothers (by 2 percentage points) and less likely to work part-time (by 0.6 percentage points) or full-time (by 1.5 percentage points).

Table 5: Marginal effects for working hours equation for mothers

	Mothers youngest child 0 to 2	Mothers youngest child 3 to 6	Mothers youngest child 7 to 15
Part-time	0.009**	–0.001	–0.006**
	(0.004)	(0.003)	(0.003)
Full-time	0.018**	–0.003	–0.015**
	(0.007)	(0.008)	(0.007)
Overtime	–0.027**	0.005	0.021**
	(0.011)	(0.011)	(0.010)

Notes: Marginal effects at mean, based on the specification of the coefficients presented in Table A2 in the Appendix (columns 2 and 4). Values in the table represent the conditional difference in the probability of working in one of the working hours options, compared to non-mothers (for mothers) and non-fathers (for fathers). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

This result, together with the descriptive statistics presented in Figure 4 which suggest that overtime in Serbia is much more frequent than part-time work, indicates that the true choice of working hours in Serbia is between full-time and overtime, rather than part-time vs. full-time. In other words, when choosing working hours (to the extent at which this is a choice), Serbian women decide if they can work overtime and if they need to. Our results suggest that when facing increased responsibility to take care of small children, women are less likely to work overtime and more likely to work full-time (or part-time). On the contrary, women who have older children, whose care can partially be transmitted to elementary schools and who generally require less attention, can use this time to work overtime to provide for the increased financial burden on their family.

6.3 Wage penalty

Table A.3 in the Appendix presents the results of the hourly wage equation. We first discuss the effects of the covariates and then turn our attention to the effect of parenthood. In line with expectations, hourly wages are higher for persons with higher levels of education and those living in urban areas and Belgrade. Additionally, married men and women have higher wages, while wages increase linearly with age for both genders. Regarding job characteristics, top occupations such as managers or professionals are associated with higher wages, while wages are higher in industry and services than in agriculture. Hourly wages are higher in

the public sector, in supervisory positions, in large firms, and in part-time work, while being employed in informal employment or as a temporary worker decreases the hourly wage.

Table 6 reveals that the motherhood penalty in the terms of hourly wages is negative, but it does not reach statistical significance (column 1).[14] However, when this effect is split by the age of the youngest child, we find that only mothers of the youngest children suffer a statistically significant penalty in wages of about 4.4% (column 2). The effect for fathers is non-significant (columns 3 and 4), indicating that fathers and non-fathers, conditional on other characteristics, have the same average levels of hourly wage. [15]

[14] Percentage interpretation of the coefficient due to the fact that, as is customary in the literature, we use the natural logarithm of the hourly wage as the dependent variable in the wage equation (rather than levels) in order to stabilize the variance of the hourly wage variable and to account for the asymmetry in the distribution of this variable. Therefore, the estimated coefficients in the wage equation, presented in Table 6 and Table A3 in the Appendix, represent the conditional change in log wages when the independent variable changes by 1, which approximately correspond to the percentage change in wages.

[15] Due to the fact that all the variables in the participation equation are also in the wage equation, the exclusion restriction condition (at least one variable has to appear in the participation equation that is not in the hours equation) for the application of the selection equation is not fulfilled. In this case it is more reasonable to adopt a model without correction, as suggested by Puhani (2000). However, the estimates in Table A.3 suggest some of the variables that enter the model are not significant (dummy variables representing number of children), but are significant in the participation equation. Therefore, dropping the insignificant variables from the wage equation enables us to fulfill the exclusion restriction condition. By using this approach, we estimate the wage equation with the selection correction. The results from this model (available upon request) are very similar to the ones presented here and reaffirm our conclusions. However, as the insignificant variables are still theoretically associated with hours of work (particularly in our framework), we opt to present the results without selection as our main results.

Table 6: Hourly wage equation

	Women		Men	
	(1)	(2)	(3)	(4)
Youngest child 0 to 15	–0.014		–0.008	
	(0.012)		(0.017)	
Youngest child 0 to 2		–0.030**		0.004
		(0.015)		(0.018)
Youngest child 3 to 6		–0.006		–0.021
		(0.015)		(0.019)
Youngest child 7 to 15		–0.011		–0.011
		(0.013)		(0.020)
Individual level controls	X	X	X	X
Job characteristics	X	X	X	X
Regional FE	X	X	X	X
Survey year FE	X	X	X	X
Observations	6,401	6,401	8,123	8,123

Notes: Results from ordinary least squares model. Individual-level controls: age, indicators for educational level, indicator for marriage status, number of adults in household, number of children in household, indicator for urban setting. Job characteristic controls: occupation, sector, ownership, informal employment, type of contract, supervising position, and firm size. Regional fixed effects: Belgrade, Vojvodina, Šumadija and Western Serbia, and Eastern and Southern Serbia. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The wage penalty for mothers with younger children indicates that if these women do not have secure jobs to return to from maternity leave, they have difficulty finding equally paid jobs that can be aligned with the cumbersome responsibility of taking care of a child. However, we do not find evidence that they continue to suffer this penalty (compared to non-mothers) in the future, as their children become older.

6.4 Robustness of results

The main results in the Estimation Results section are estimated using the probit model (participation rate), with ordered probit (hours worked per week) or with ordinary least squares (log hourly wage). Tables 1 and 2 show that both mothers and non-mothers, and fathers and non-fathers, have significantly different background characteristics. In this section we aim to account for these differences by

applying matching methodology. We estimate the parenthood penalty, using propensity score matching to examine the robustness of the findings.

The key aim of propensity score matching is achieving balanced observable characteristics by weighting observations differently when estimating the treatment effect (in our case, parenthood). This propensity score matching requires satisfying the conditional independence assumption (CIA): the selection into treatment (parenthood) is based solely on observable characteristics. The second assumption is common support, which ensures that female/males with the same characteristics have a positive probability of being treated.

In a first stage we estimate the propensity score variable $p(X)$ using a probit model of being a mother/father using the observable characteristics. The estimator of the average treatment of the treated using the propensity score $p(X_i)$ is then given by Rosenbaum and Rubin (1983):

$$\tau^{PSM} = E\Big[E\{Y_{1i}|D_i=1, p(X_i)\} - E\{Y_{0i}|D_i=0, p(X_i)\}|D_i=1\Big]$$

where the outer distribution is over ($p(X_i)\,|\,D_i=1$) and Y_{1i} and Y_{0i} are potential outcomes in the two counterfactual situations of treatment and no treatment (parent and no parent, in our case). Stated differently, the propensity score estimator is the mean difference in outcomes of the parents and non-parents over the common support and weighted by the propensity score distribution of participants.

We now estimate again the labour market outcomes (participation rate, hours worked per week, and log hourly wage) for parents and non-parents, but in this section we use propensity score matching. Standard propensity score matching can only be applied to estimate the results for mothers/fathers as one group and it is not possible to disaggregate parents by the age of the youngest child.

The results are shown in Table 7 and confirm our main findings. We find that the participation of fathers is higher than the participation of non-fathers, while the participation of mothers overall does not differ from the participation of non-

mothers. We do not find any impact on hours worked per week and hourly wage of parents.

Table 7: Robustness: Labour market outcome of parents and non-parents

	Participation		Hours worked per week		Hourly wage equation	
	Women	Men	Women	Men	Women	Men
	(1)	(2)	(3)	(4)	(5)	(6)
Child 0 to 15	–0.020	0.027***	0.035	–0.008	–0.005	–0.046
	(0.008)	(0.011)	(0.021)	(0.021)	(0.029)	(0.025)
Individual-level controls	X	X	X	X	X	X
Regional FE	X	X	X	X	X	X
Survey year FE	X	X	X	X	X	X
Observations	29,669	35,311	13,243	17,307	6,401	8,123

Notes: Marginal effects from probit model. Individual level controls: age, age squared, indicators for educational level, indicator for marriage status, number of adults in household, indicator for urban setting. Regional fixed effects: Belgrade, Vojvodina, Šumadija and Western Serbia, and Eastern and Southern Serbia. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

7. DISCUSSION AND CONCLUSION

Women in Serbia have worse labour market outcomes than men and one channel that can explain these differences between genders is motherhood. The arrival of a child in the household brings additional responsibilities for the parents, and women usually take on more child-rearing responsibilities. This implies that not only are women detached from the labour market during parental leave, but their household responsibilities increase permanently with the arrival of a child. The difficulty of reconciling parenting responsibilities and general chores can overburden women and make it more difficult for them to compete in the labour market. The phenomenon of women's position in the labor market changing when they become mothers is called the motherhood penalty, and we explore this phenomenon, together with a similar analysis for fathers, in the Serbian context.

We find that on average there is no difference between mothers and non-mothers in terms of labour market participation. However, women with younger children

are less likely to be active in the labour market than both women with older children and non-mothers. This pattern is not found for men: to the contrary, all fathers, regardless of the age of children, have a higher propensity to be active in the labour market than non-fathers. These findings together show that a traditional 'breadwinner' model prevails in Serbia, where women are predominantly providers of childcare and stay at home, while men assume the role of primary breadwinner and increase their activity to compensate for the lower labour participation of mothers. However, as the children grow older they require less care and the burden of childcare shifts towards elementary schools, and women can return to the labour market. Indeed, women whose youngest child is aged 7 to 15 have the highest participation rate among the women in our sample. One explanation for women with older children having a higher propensity to be active than non-mothers could be that households with children have larger expenses, and when available for work mothers become active in the labour market so that they contribute to the household budget.

This finding is also reflected in the working hours of employed mothers and fathers. Contrary to the findings for countries where part-time work for women is common, in Serbia part-time work options are limited, and main distinction is between working full-time or overtime. Mothers of younger children are less likely to work overtime than non-mothers, while mothers of older children are more likely to work overtime than childless women. Similar explanations apply: young mothers need to restrict their working hours to cope with household chores, while mothers of older children have more expenses and this is reflected in the longer hours worked.

Finally, we do not find significant differences between mothers and non-mothers in terms of hourly wages. However, when mothers are grouped according to the age of the youngest child, a difference emerges between the wages of mothers and non-mothers, and in particular mothers with a very young child (aged 0 to 2). One explanation of this effect could be the difficulty women face finding jobs after maternity leave that pay the same as for women with the same characteristics and can be aligned with the cumbersome responsibilities of taking care of an infant. We do find that mothers of older children (3 to 15 years) earn slightly less than non-mothers, but this difference does not reach statistical significance. No difference in hourly wages is found for fathers and non-fathers.

To summarize our results, we find that, conditional on observables, mothers of younger children (0 to 2 years) have lower activity rates, are less likely to work overtime, and have lower hourly wage rates than non-mothers, but this difference disappears as the children get older. In fact, women with children aged 7 to 15 are more likely to be active in the labour market regarding both the extensive margin (participation in the labour market) and the intensive margin (increased working hours). Overall, in the Serbian setting, motherhood does not seem to impact the labour market trajectories of women in the long run. Although observed effects in early parenthood can account for some gender differences in labour market outcomes, other factors such as the traditional role of women in the household, labour market discrimination, and the difficult position of women in rural areas are at play. One limitation of our study is that we exclude from our sample mothers and fathers who fall into the age group 51 years or older and who have children 15 years old or younger. Our findings do not speak to this group of parents. Additionally, our sample does not include divorced fathers who do not live with their child(ren), which could potentially bias the results of the fathers' outcomes.

In some aspects the case of Serbia fits well with other results available for post-communist countries (Cukrowska-Torzewska and Matysiak, 2020; Lebedinski et al., 2020). The lack of the part-time options and low employment flexibility typically means that after an initial period of taking care of infants (children aged 0-2 years) women in Serbia return to the work they previously had, with the same working hours and wage rates. However, in contrast to findings for other post-communist countries, the penalties in Serbia do not seem to last long. Faced with the increased financial burden on the household and the generally low wages and living standards in Serbia, women are expected to contribute to the household budget, while at the same time performing most of the household and childcare duties.

Two policies should be considered based on our findings. First, increasing childcare availability for the age group 1 to 2 years could help alleviate the domestic burden and enable women to return to the labour market, as this group of women is least likely to be active. Second, paternity leave policies that allocate part of parental leave exclusively to fathers should be explored. Paternity leave policies are a prominent way to incentivize fathers to take more responsibility for childcare

and this policy could help increase fathers' involvement in the household and distribute housework more equally (Schober, 2014; Patnaik, 2019; Farré and Gonzalez, 2019).

Acknowledgment

We would like to thank the Statistical Office of the Republic of Serbia for granting us access to the individual-level datasets of the Labour Force Survey for the period 2014 to 2018.

REFERENCES

Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labour Economics*, *34*(3), 545–579. https://doi.org/10.1086/684851

Anić, A., & Krstić, G. (2019). What lies behind the gender wage gap in Serbia? *Economic Annals*, *64*(223), 137–169. https://doi.org/10.2298/EKA1923137A

Avlijaš, S., Nevena I., Vladisavljević, M., and Vujić, S. 2013. *Gender Pay Gap in the Western Balkan Countries: Evidence from Serbia, Montenegro and Macedonia*. Belgrade: Foundation for the Advancement of Economics. http://ebooks.ien.bg.ac.rs/579/1/Gender%20pay%20gap%20in%20the%20Western%20balkan%20countries.pdf

Arandarenko, M., Vladisavljević, M., & Žarković-Rakić, J. (2012). From inactivity to work: analysis of the public policies impact and factors affecting inactivity. *In* Cerović et al. ed. *From Global Crisis to Economic Growth. Which Way to Take*? Volume I, Economics. Faculty of Economics, University of Belgrade, 2012.

Babović, M. (2018). Gender Equality Index 2018: Measuring gender equality in Serbia 2016. *Belgrade: Social Inclusion and Poverty Reduction Unit Government of Republic of Serbia. http://socijalnoukljucivanje.gov.rs/wpcontent/uploads/2018/12/Indeks_rodne_ravnopravnosti_u_Republici_Srbiji_2018_eng.pdf*

Baker, M., & Milligan, K. (2008). How does job-protected maternity leave affect mothers' employment? *Journal of Labour Economics*, *26*(4), 655–691. https://doi.org/10.1086/591955

Becker, G. (1981). A Treatise on the Family. *Cambridge, MA, Harvard University Press*.

Blau, F. D., & Kahn, L. M. (2013). Female labour supply: Why is the United States falling behind? *American Economic Review*, *103*(3), 251–56. http://dx.*doi*.org/10.1257/aer.103.3.251

Budig, M. J., Misra, J., & Boeckmann, I. (2012). The motherhood penalty in cross-national perspective: The importance of work–family policies and cultural attitudes. *Social Politics*, *19*(2), 163–193. https://doi.org/10.1093/sp/jxs006

Budig, M. J., Misra, J., & Boeckmann, I. (2016). Work–family policy trade-offs for mothers? Unpacking the cross-national variation in motherhood earnings penalties. *Work and Occupations*, *43*(2), 119–177. https://doi.org/10.1177/0730888415615385

Cukrowska-Torzewska, E., & Matysiak, A. (2020). The motherhood wage penalty: A meta-analysis. *Social Science Research*, *88*, 102416. https://doi.org/10.1016/j.ssresearch.2020.102416

Farré, L., & González, L. (2019). Does paternity leave reduce fertility? *Journal of Public Economics*, *172*, 52–66. https://doi.org/10.1016/j.jpubeco.2018.12.002

Eurostat (2019). *How do Women and Men Use Their Time - Statistics. https://bit.ly/3lKxEZ9. Accessed: Nov. 18th 2020.*

Eurostat (2021). *Estimated Average Age of Young People Leaving the Parental Household by Sex*. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=yth_demo_030&lang=en

Givord, P., & Marbot, C. (2015). Does the cost of child care affect female labour market participation? An evaluation of a French reform of child care subsidies. *Labour Economics*, *36*, 99–111. https://doi.org/10.1016/j.labeco.2015.07.003

Gray, J. S. (1997). The fall in men's return to marriage: Declining productivity effects or changing selection? *Journal of Human Resources*, 481–504. https://doi.org/10.2307/146180

Grimshaw, D., & Rubery, J. (2015). *The Motherhood Pay Gap*. Geneva: International Labour Organization.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.

Hodges, M. J., & Budig, M. J. (2010). Who gets the daddy bonus? Organizational hegemonic masculinity and the impact of fatherhood on earnings. *Gender & Society*, *24*(6), 717–745. https://doi.org/10.1177/0891243210386729

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019). Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*, 109, 22–26. https://*doi*.org/10.1257/pandp.20191078

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2020). *Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation* (No. w28082). National Bureau of Economic Research.

Lebedinski, L., Perugini, C., & Vladisavljevic, M. (2020). Child Penalty in Russia: Evidence from an Event Study. IZA Discussion Paper No. 13928

Milić, J., & Zhou, J. (2015). Youth Residential Situation in the Transition to Adulthood: The Case of Serbia (An Exploratory Approach). *Stanovništvo*, *53*(1), 61–87. https://doi.org/10.2298/STNV1501061M

OECD (2012). *Closing the Gender Gap: Act Now*. Paris: Organisation for Economic Cooperation and Development.

OECD (2020), *Education at a Glance 2020: OECD Indicators*. OECD Publishing, Paris, https://doi.org/10.1787/69096873-en.

Patnaik, A. (2019). Reserving time for daddy: The consequences of fathers' quotas. *Journal of Labour Economics*, *37*(4), 1009–1059. https://doi.org/10.1086/703115

Puhani, P. (2000). The Heckman Correction for sample selection and its critique. *Journal of Economic Surveys*. **14** (1). https://*doi*.org/10.1111/1467-6419.00104

Schober, P. S. (2014). Parental leave and domestic work of mothers and fathers: A longitudinal study of two reforms in West Germany. *Journal of Social Policy*, *43*(2), 351. https://doi.org/10.1017/S0047279413000809

Schönberg, U., & Ludsteck, J. (2014). Expansions in maternity leave coverage and mothers' labour market outcomes after childbirth. *Journal of Labour Economics*, *32*(3), 469–505. https://doi.org/10.1086/675078

Simonsen, M. (2010). Price of high-quality daycare and female employment. *Scandinavian Journal of Economics*, *112*(3), 570–594. https://doi.org/10.1111/j.1467-9442.2010.01617.x

Smith Koslowski, A. (2011). Working fathers in Europe: Earning and caring. *European Sociological Review*, *27*(2), 230–245. https://doi.org/10.1093/esr/jcq004

Statistical office of the Republic of Serbia (2019a). Demographic Yearbook 2002–2003.

Statistical office of the Republic of Serbia (2017). Labour Force Survey – Methodological instruction.

Statistical office of the Republic of Serbia (2019a). Demographic Yearbook 2018.

Statistical office of the Republic of Serbia (2019b). Preschool education questionnaire 2018.

Statistical office of the Republic of Serbia, SORS (2020a). Labour Force Survey in the Republic of Serbia, 2019. Bulletin no 658

Statistical office of the Republic of Serbia, SORS (2020b). DevInfo profile: Republic of Serbia.

Statistical office of the Republic of Serbia, SORS (2021). Children in preschool education in Serbia, 2011–2020. Available at https://publikacije.stat.gov.rs/G2021/Pdf/G202127005.pdf (accessed on 2 November 2021)

Vladisavljević, M., Avlijaš, S. and Vujić, S. (2015). Gender inequality in wages in the Western Balkans. *In* C. Perugini and F. Pompei (Eds.), *Inequalities During and After Transition in Central and Eastern Europe,* pp. 222–243. Basingstoke: Palgrave Macmillan UK. https://doi.org/10.1057/9781137460981

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Ch. 17. Cambridge, MA, MIT Press. https://doi.org/10.1007/s00712-003-0589-6

Žarkovic-Rakic, J., and Vladisavljevic, M. 2016. *Women's Access To Economic Opportunities in Serbia*. World Bank Report, No. 96451-Yf. Washington, DC: The World Bank Group. https://doi.org/10.1596/25183

Received: June 18, 2021
Accepted: December 22, 2021

APPENDIX

Figure A1: Usual hours worked per week, by gender and parenthood

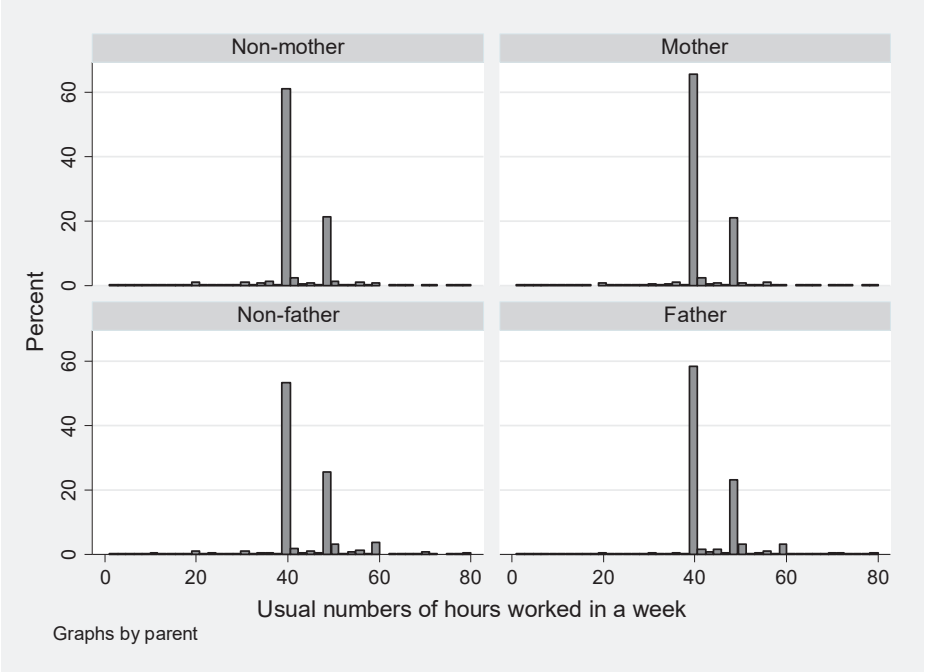


Table A1: Full estimation of the participation penalty

	(1)	(2)	(3)	(4)
	Women		Men	
Youngest child 0 to 15	0.002		0.067***	
	(0.008)		(0.008)	
Youngest child 0 to 2		–0.075***		0.070***
		(0.010)		(0.010)
Youngest child 3 to 6		0.004		0.063***
		(0.010)		(0.010)
Youngest child 7 to 15		0.063***		0.068***
		(0.009)		(0.009)
Age	0.090***	0.083***	0.051***	0.051***
	(0.006)	(0.006)	(0.004)	(0.004)

Age squared	–0.001***	–0.001***	–0.001***	–0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
General or VET secondary	0.169***	0.171***	0.088***	0.088***
	(0.008)	(0.008)	(0.005)	(0.005)
College/university or higher	0.330***	0.341***	0.147***	0.147***
	(0.009)	(0.009)	(0.007)	(0.007)
Married	–0.011	–0.002	0.055***	0.055***
	(0.007)	(0.007)	(0.006)	(0.006)
Number of adults in household	–0.003	0.000	–0.003	–0.003
	(0.003)	(0.003)	(0.002)	(0.002)
2 children	–0.010	–0.012	–0.014*	–0.013
	(0.008)	(0.008)	(0.008)	(0.008)
3 or more children	–0.070***	–0.050***	–0.041***	–0.041***
	(0.011)	(0.011)	(0.011)	(0.011)
Urban	0.001	0.004	–0.036***	–0.036***
	(0.006)	(0.006)	(0.004)	(0.004)
Vojvodina	–0.020**	–0.022***	–0.007	–0.007
	(0.008)	(0.008)	(0.006)	(0.006)
Šumadija and Western Serbia	–0.004	–0.008	–0.001	–0.001
	(0.008)	(0.008)	(0.006)	(0.006)
Eastern and Southern Serbia	–0.032***	–0.039***	–0.022***	–0.022***
	(0.008)	(0.008)	(0.006)	(0.006)
Year 2015	–0.009	–0.008	–0.007	–0.007
	(0.009)	(0.009)	(0.007)	(0.007)
Year 2016	0.004	0.004	–0.002	–0.002
	(0.009)	(0.009)	(0.006)	(0.006)
Year 2017	0.009	0.010	0.002	0.002
	(0.009)	(0.009)	(0.006)	(0.006)
Year 2018	0.018**	0.020**	0.015**	0.015**
	(0.009)	(0.008)	(0.006)	(0.006)
Observations	21,335	21,335	25,868	25,868

Table A2: Full estimation of the hours penalty

	(1)	(2)	(3)	(4)
	Women		Men	
Youngest child 0 to 15	0.022		0.019	
	(0.032)		(0.033)	
Youngest child 0 to 2		–0.098**		–0.012
		(0.041)		(0.039)
Youngest child 3 to 6		0.018		0.032
		(0.040)		(0.039)
Youngest child 7 to 15		0.076**		0.039
		(0.035)		(0.038)
Age	–0.002	–0.004**	–0.004***	–0.005***
	(0.002)	(0.002)	(0.001)	(0.001)
Primary or less (omitted)				
General or VET secondary	–0.004	–0.006	–0.133***	–0.133***
	(0.064)	(0.064)	(0.041)	(0.041)
College/university or higher	–0.187***	–0.178**	–0.313***	–0.311***
	(0.070)	(0.070)	(0.048)	(0.048)
Married	–0.109***	–0.093***	–0.010	–0.009
	(0.030)	(0.030)	(0.030)	(0.030)
Number of adults in household	0.043***	0.045***	0.030***	0.030***
	(0.011)	(0.011)	(0.009)	(0.009)
one child (omitted)				
two children	–0.006	–0.008	0.022	0.018
	(0.031)	(0.031)	(0.029)	(0.029)
three or more children	–0.091	–0.057	0.003	0.007
	(0.057)	(0.057)	(0.049)	(0.050)
Urban	–0.101***	–0.098***	–0.094***	–0.093***
	(0.027)	(0.027)	(0.021)	(0.021)
Managers (omitted)				
Professional	–0.227***	–0.226***	–0.062	–0.061
	(0.081)	(0.081)	(0.062)	(0.062)
Technicians and ass. prof.	–0.058	–0.055	0.065	0.066
	(0.083)	(0.083)	(0.065)	(0.065)
Clerical support workers	–0.124	–0.121	0.174**	0.175**
	(0.085)	(0.085)	(0.069)	(0.069)
Service and sales workers	0.741***	0.742***	0.627***	0.627***
	(0.087)	(0.087)	(0.067)	(0.068)
Skilled agricultural workers	–0.486	–0.455	0.242	0.243
	(0.409)	(0.406)	(0.167)	(0.167)
Craft and trade workers	0.447***	0.447***	0.383***	0.383***
	(0.099)	(0.099)	(0.068)	(0.068)

Plant and machine operators	0.089	0.088	0.449***	0.449***
	(0.099)	(0.099)	(0.068)	(0.068)
Elementary occupations	0.022	0.022	0.286***	0.287***
	(0.096)	(0.096)	(0.074)	(0.074)
Agriculture (omitted)				
Industry	−0.082	−0.083	0.225***	0.225***
	(0.159)	(0.159)	(0.074)	(0.074)
Services	−0.125	−0.126	0.167**	0.166**
	(0.158)	(0.158)	(0.075)	(0.075)
Public sector	−0.490***	−0.494***	−0.511***	−0.511***
	(0.028)	(0.028)	(0.020)	(0.020)
Informally employed	−0.815***	−0.821***	−0.486***	−0.486***
	(0.067)	(0.067)	(0.043)	(0.043)
supervising position	0.054	0.051	0.088***	0.087***
	(0.034)	(0.034)	(0.027)	(0.027)
temporary worker	−0.176***	−0.183***	−0.102***	−0.103***
	(0.030)	(0.030)	(0.023)	(0.023)
10 employees or less				
11 to 49 employees	−0.013	−0.012	0.030	0.030
	(0.034)	(0.034)	(0.028)	(0.028)
50 employees or more	−0.017	−0.017	−0.016	−0.016
	(0.031)	(0.031)	(0.024)	(0.024)
Belgrade (omitted)				
Vojvodina	−0.056*	−0.057*	−0.100***	−0.100***
	(0.031)	(0.031)	(0.026)	(0.026)
Šumadija and Western Serbia	0.190***	0.187***	0.063**	0.063**
	(0.032)	(0.032)	(0.027)	(0.027)
Eastern and Southern Serbia	0.038	0.033	−0.026	−0.027
	(0.033)	(0.033)	(0.027)	(0.027)
Year 2014 (omitted)				
Year 2015	0.049	0.049	−0.001	−0.001
	(0.040)	(0.040)	(0.033)	(0.033)
Year 2016	0.014	0.013	−0.062*	−0.061*
	(0.038)	(0.038)	(0.032)	(0.032)
Year 2017	0.013	0.014	−0.006	−0.006
	(0.038)	(0.038)	(0.032)	(0.032)
Year 2018	0.021	0.022	−0.049*	−0.048*
	(0.035)	(0.035)	(0.029)	(0.029)
Constant cut1	−2.338***	−2.405***	−1.997***	−2.011***
	(0.206)	(0.207)	(0.124)	(0.124)
Constant cut2	0.380*	0.315	0.298**	0.284**
	(0.205)	(0.205)	(0.122)	(0.122)
Observations	13,243	13,243	17,307	17,307

Table A3: Full estimation of the wage penalty

	(1)	(2)	(3)	(4)
	Women		Men	
Youngest child 0 to 15	–0.014		–0.008	
	(0.012)		(0.017)	
Youngest child 0 to 2		–0.030**		0.004
		(0.015)		(0.018)
Youngest child 3 to 6		–0.006		–0.021
		(0.015)		(0.019)
Youngest child 7 to 15		–0.011		–0.011
		(0.013)		(0.020)
Age	0.003***	0.003***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Primary or less (omitted)				
General or VET secondary	0.099***	0.099***	0.105***	0.105***
	(0.019)	(0.019)	(0.016)	(0.016)
College/university of higher	0.228***	0.229***	0.271***	0.270***
	(0.022)	(0.022)	(0.020)	(0.020)
Married	0.018*	0.020*	0.077***	0.077***
	(0.011)	(0.011)	(0.015)	(0.015)
Number of adults in household	–0.022***	–0.021***	–0.022***	–0.022***
	(0.004)	(0.004)	(0.004)	(0.004)
one child (omitted)				
two children	0.015	0.015	–0.015	–0.013
	(0.011)	(0.011)	(0.014)	(0.014)
three or more children	0.015	0.018	0.014	0.014
	(0.022)	(0.023)	(0.022)	(0.022)
Urban	0.042***	0.043***	0.019**	0.019**
	(0.009)	(0.009)	(0.009)	(0.009)
Managers (omitted)				
Professional	–0.157***	–0.155***	–0.070	–0.071
	(0.052)	(0.052)	(0.047)	(0.047)
Technicians and ass prof.	–0.314***	–0.312***	–0.160***	–0.161***
	(0.052)	(0.052)	(0.047)	(0.047)
Clerical support workers	–0.376***	–0.374***	–0.290***	–0.291***
	(0.053)	(0.053)	(0.048)	(0.048)
Service and sales workers	–0.567***	–0.565***	–0.346***	–0.347***
	(0.053)	(0.053)	(0.047)	(0.047)
Skilled agricultural workers	–0.323**	–0.317**	–0.346***	–0.346***
	(0.145)	(0.145)	(0.097)	(0.097)
Craft and trades workers	–0.600***	–0.598***	–0.296***	–0.296***
	(0.055)	(0.055)	(0.048)	(0.048)

Plant and machine operators	–0.495***	–0.493***	–0.277***	–0.278***
	(0.055)	(0.055)	(0.047)	(0.047)
Elementary occupations	–0.592***	–0.591***	–0.399***	–0.399***
	(0.053)	(0.054)	(0.049)	(0.049)
Agriculture (omitted)				
Industry	0.204**	0.204**	0.149***	0.149***
	(0.083)	(0.083)	(0.029)	(0.029)
Services	0.151*	0.151*	0.082***	0.082***
	(0.083)	(0.083)	(0.029)	(0.029)
Public sector	0.075***	0.075***	0.142***	0.142***
	(0.014)	(0.014)	(0.010)	(0.010)
Informally employed	–0.117***	–0.118***	–0.129***	–0.129***
	(0.025)	(0.025)	(0.024)	(0.024)
supervising position	0.115***	0.115***	0.164***	0.164***
	(0.018)	(0.018)	(0.015)	(0.015)
temporary worker	–0.102***	–0.103***	–0.090***	–0.090***
	(0.011)	(0.011)	(0.010)	(0.010)
part–time worker	0.230***	0.230***	0.128***	0.128***
	(0.035)	(0.035)	(0.044)	(0.044)
10 employees or less				
11 to 49 employees	0.061***	0.061***	0.073***	0.074***
	(0.013)	(0.013)	(0.013)	(0.013)
50 employees or more	0.118***	0.118***	0.120***	0.120***
	(0.013)	(0.013)	(0.011)	(0.011)
Belgrade (omitted)				
Vojvodina	–0.090***	–0.090***	–0.112***	–0.112***
	(0.012)	(0.012)	(0.013)	(0.013)
Šumadija and Western Serbia	–0.140***	–0.140***	–0.179***	–0.179***
	(0.012)	(0.012)	(0.013)	(0.013)
Eastern and Southern Serbia	–0.170***	–0.171***	–0.192***	–0.192***
	(0.012)	(0.012)	(0.013)	(0.013)
Year 2014 (omitted)				
Year 2015	0.003	0.002	-0.001	-0.001
	(0.016)	(0.016)	(0.016)	(0.016)
Year 2016	0.021	0.021	0.020	0.020
	(0.015)	(0.015)	(0.015)	(0.015)
Year 2017	0.021	0.021	0.031**	0.031**
	(0.015)	(0.015)	(0.015)	(0.015)
Year 2018	0.064***	0.064***	0.088***	0.088***
	(0.014)	(0.014)	(0.014)	(0.014)
Constant	5.102***	5.106***	5.041***	5.038***
	(0.103)	(0.103)	(0.063)	(0.063)
Observations	6,401	6,401	8,123	8,123
R-squared	0.475	0.475	0.343	0.344

ECONOMIC ANNALS, Volume LXVII, No. 232 / January – March 2022
UDC: 3.33 ISSN: 0013-3264

https://doi.org/10.2298/EKA2232049D

*Uros Delevic**
*James Kennell***

MULTINATIONALS AND WAGES: EVIDENCE FROM EMPLOYER–EMPLOYEE DATA IN SERBIA

ABSTRACT: *Many studies have reported that foreign-owned companies pay higher wages on average than domestic companies. However, this can be attributed to the different composition of the workforce or to a wage premium at the individual worker level. This paper contributes to this literature by observing whether individuals that change their job from a domestic to a foreign-owned company experience a change in their wages. Furthermore, it investigates whether this difference in wage patterns is moderated by workers' education. This paper is one of the very few micro-econometric studies that deal with this question in a transition country, Serbia, using employer–employee data on the private sector over a long time period (15 years). Changing jobs is found to be positively associated with workers' wages: the change in wages is higher when moving from a domestic to a foreign company than vice versa. The evidence suggests that more-educated workers benefit the most from leaving domestic companies.*

KEY WORDS: *labour mobility, job change, wage change, MNEs and wages, transition economies*

JEL CLASSIFICATION: B21, C12, C01, C33, F16, F66

* University of Greenwich, u.delevic@gre.ac.uk
** University of Greenwich, j.s.kennell@gre.ac.uk

1. INTRODUCTION

This paper is about wage dynamics related to changing jobs. Specifically, we analyse labour mobility and wage change. This paper does not address the determinants of labour mobility per se – why workers are more likely to move from a domestic to a foreign company or vice versa – but investigates the effect of labour mobility on wages.

The investigation of the difference between wages paid by foreign and domestic companies, known as the 'foreign-wage premium', is a longstanding area of research. Various studies have found that multinational enterprises (MNE) pay higher wages than domestic companies (for example, Aitken et al., 1996; Lipsey and Sjöholm, 2004), although this claim has been contested recently and has produced calls for studies that reconsider this generic claim, particularly in the context of developing countries (Coniglio et al., 2015; van der Straaten et al., 2020; Cruz et al., 2018). Additionally, whether there is a true company ownership effect or whether the foreign wage premium is the result of workers' different characteristsics is a matter for discussion. In other words, multinational companies may on average pay higher wages than domestic companies because they employ more-skilled workers, who are normally paid higher wages (Doh, 2019). In developed countries the competitive advantage of multinational enterprises and domestic companies is similar, but in developing countries the foreign wage premium is likely to be higher (Glass and Saggi, 2002) because of the higher quality of jobs created by MNEs compared with domestic companies due to technological superiority and managerial sophistication (Gereffi et al., 2019; Pandit et al., 2020).

Labour mobility is a key channel through which MNEs can affect host country development (Masso & Vahter, 2019). Research has identified that workers moving from MNEs to local firms bring specific new or enhanced skills and experience with them (Davis & Poole, 2020), spreading new skills and training to the domestic sector (Poole, 2013; Pradhan, 2006). Additionally, studies have analysed the creation of 'spinoffs', new firms established by former MNE employees in the same industry (Andersson and Klepper, 2013).

In these cases, the research highlights how MNEs foster local development by increasing the potential for knowledge transfer (Park et al., 2021). The focus of

the current study, however, is on the relationship between MNEs and their host economies in terms of workers' welfare, which could lead to economic growth by increasing aggregate demand (Woodgate, 2021). This study considers this relationship using a novel data source and attempts to control for the influence of worker characteristics. As Earle et al. (2012) explain, studies with firm-level data are usually missing information about worker characteristics, so it is not possible to control for the effect of worker heterogeneity on the foreign wage premium. To the extent that a foreign wage premium exists, entry or expansion of foreign MNEs can benefit the workers who are employed in these companies (Glass and Saggi, 2002). Foreign companies not only contribute to employment but also train labour, which may become available to local competitors or result in the establishment of new businesses (Dunning, 1988; Narula, 1996). However, this may also lead to the crowding-out of domestic companies, which then face stiffer competition for a limited pool of skilled workers and face higher costs to attract talented workers (Ayentemi et al., 2018).

In order to develop an understanding of whether multinational companies – specifically foreign-owned firms operating in a given host country – pay higher wages than domestic companies for similar workers, it is crucial to control for workers' characteristics. In this study, worker characteristics are kept constant while observing whether changing jobs from domestic firms to MNEs affects workers' wages. Thus, this research analyses whether the higher wages paid by MNEs are due to the different characteristics of workers.

This study uses linked employer–employee data from Serbia for 2000 to 2014, provided by the Serbian Social Register. This is the first study to use data from this source, which is held by the national pension agency, part of the Ministry of Employment and Social Welfare. The object of this paper is to discover whether workers who move from domestic to foreign-owned companies experience a larger wage increase than those who do not change jobs (or move from foreign to domestic-owned companies), and whether the change in wages is moderated by the level of workers' education.

This research finds not only that MNEs in Serbia pay higher wages on average, but also that they pay higher wages to workers with similar characteristics, so the wage change is the result of changing from a domestic to a foreign company. The

wage increase is higher for more-educated workers, who are more able to benefit from the entry of MNEs to Serbia. This effect is significant, with the wages of more-educated workers who move from domestics to foreign companies increasing by 21%.

The remainder of this paper is organised as follows. Section 2 presents the literature review. Section 3 presents the methodology. The results are given in Section 4, and section 5 provides concluding remarks and some policy recommendations.

2. LITERATURE REVIEW

This literature review is organised around the three main dimensions of the paper: (1) the effect of general labour mobility on wages, regardless of the type of company; (2) labour mobility and the effect of company ownership on wage change; and (3) the moderating effect of worker-specific characteristics, including education, on wage change.

2.1 Labour mobility and wage change

Early research on job changes took a transactional, cost-benefit approach to understanding why people change jobs, and the impact this has on the employee. According to Kidd (1991), there are two forces that underly changing jobs. On the one hand, labour mobility benefits employees, as it rewards time spent on job search and training. In an analysis of wage growth and job turnover, Bartel and Borjas (1981) find that labour mobility substantially determines not only wage levels but also the dynamics of wage growth. However, this pattern is also age-dependent. Workers who change jobs voluntarily experience a wage increase if they change early in their career, but labour mobility negatively affects the wages of more-senior workers who decide to change jobs later in their career. Abbott and Beach (1994) investigate the dynamics of wage change resulting from female workers changing jobs in Canada and find that changing jobs results in short-run wage increases of up to 9%. From this perspective, labour turnover is the result of expected positive returns from investing time and resources in labour mobility (Kid, 1991). On the other hand, employees who change jobs can also experience negative impacts if they forego job-specific skills accumulated with their previous employer. Abbott and Beach (1994) also argue that those who change jobs can

bear costs due to the loss of previous investments in training, so that the gains from changing jobs need to be higher than these sunk costs.

Identifying the difference between those whose wages change because they change company and those who experience a wage change within the same company is complex. The most accurate way to estimate short-run wage change is to compare workers who change jobs with those who stay in the same job, but in the long run this is difficult since unobservable factors can influence wage growth for stayers. Campbell (2001) points this out and finds that over a three-year period about 10% of wage increases arise from changing jobs. Only 40% of wage growth is related to job change, while the rest is related to a higher rate of wage growth. Widerstedt (1998) finds that returns on work experience are higher for workers who change jobs than for those who do not. However, this is not directly caused by the change but rather by the accumulation of knowledge and experience. In sum, previous research does not provide a conclusive answer as to whether changing jobs positively affects wages. It is suggested that this is an empirical question whose answer may depend on several contextual characteristics. Notwithstanding this lack of consensus on the relationship between changing jobs and wage growth, the underlying motive for changing company is an expectation that it will be followed by a positive change in either wages or job quality (Lisi, 2018), both of which lead to higher employee satisfaction, which is considered an important motive for changing jobs.

Changing employer has been shown to have a particularly positive impact on job satisfaction if it does not involve changing occupation (Zhou et al., 2017). There is an expectation that changing to a similar job will also have a positive impact on workers' wages because of their experience with their previous employer, whereas if changing jobs also means changing occupation, previous work experience might be worthless (Heinrichs et al., 2020). Zhou et al. (2017) highlight the difference between those who change both job and occupation and those who change their job within the same occupation. Both groups experience increased job satisfaction when they change, described as the 'honeymoon' effect. However, in the second year after the change the latter group experiences declining job satisfaction, which then increases slightly, while the former group experiences a decline in job satisfaction that does not increase later – the 'hangover' effect.

Longhi and Brynin (2010) combine these two approaches in an analysis of occupational change in Great Britain and Germany, and find that a change in jobs which also involves a change of occupation is, in general, beneficial for employees, as the wages of workers who change jobs increase more than the wages of workers who do not. The study focuses on wages as the main aspect of job satisfaction, which it measures at the moment of change, even though job satisfaction is normally understood to be a multi-dimensional construct (Knežević et al., 2020). Although job satisfaction cannot be reduced to its financial aspect, wage studies highlight that wages are an important element (Bossler & Brozseit, 2017; Hamermesh, 1999).

2.2 The effect of foreign ownership

Conceptual literature

The literature not only shows that changing jobs influences workers' wages but also that MNEs pay higher wages than domestic companies. Several studies have evaluated the impact of Foreign Direct Investment (FDI) on wages. A review of these studies by Barba Navarettiet al. (2004) suggests that MNEs support labour development in host countries by offering new jobs that require higher skills, thereby encouraging students to attend university (Blomstrom, 2002).

Some have argued that MNEs pay higher wages simply because they tend to be concentrated in knowledge- and technology-intensive industries that require higher wages (Sahu & Goel, 2019). Research on MNEs posits that in order to attract FDI, companies need to have firm-specific (or ownership-specific) advantages such as knowledge or technology, so MNEs will emerge in knowledge- and technology-intensive industries and create high-quality jobs that require more-educated workers (Dunning, 1998). If advanced technologies and knowledge are the main source of MNEs' advantage, foreign companies will demand more-educated workers whose compensation will be higher.

Over time, MNEs find it expensive to pay expatriates and so train local labour to take over some of the technical and managerial positions (Fosfuri et al., 2001). However, labour mobility means that MNE workers may start working for domestic companies, leading to knowledge spillover and thus increasing the competitive advantage of domestic competitors. Gorg et al. (2007), in their analysis of Ghanaian employers, find that on average employees of foreign

companies stay longer with the same employer than those in domestic companies. It has been argued that a reason for this reduced labour mobility is an increase in workplace training, because MNEs perceive investing in labour as a way to keep workers and avoid the spillover of knowledge to local competitors (Fosfuri et al. 2001), reducing the potential knowledge exchange benefits for host countries from FDI by MNEs (Poole, 2013; Pradhan, 2006). Miyamoto (2003) suggests that by investing in employee training, MNEs not only develop worker skills but also gradually increase the quality of MNE operations. The increased skill base of local labour then helps to attract better quality FDI, associated with higher-wage jobs.

Although much previous research focuses on the relationship between FDI and wages in terms of the novel job roles and higher skill requirements of MNEs, other studies examine the determinants of wages when jobs in MNEs and domestic firms are similar (Heyman, 2007). The causes of this MNE wage premium are fourfold. First, the likelihood that MNEs will close plants and offices and reallocate their activities (Bernard and Sjoholm, 2003; Dewit et al., 2019) results in lower job security, which has to be compensated for with higher wages. Once controlled for size and performance, MNEs are more footloose and more likely to close than domestic companies.

Second, MNEs may be forced to pay higher wages due to labour market information asymmetry. MNEs may be in a disadvantaged position when it comes to finding the best workers because of their lack of integration in local networks. Therefore, labour market imperfections induce MNEs to offer higher wages to attract the best workers (Dobbelaere & Kiyota, 2018; Girma and Gorg, 2007).

Third, to the extent that foreign companies are more productive than domestic companies and that their productivity advantage comes from ownership advantages such as technology or knowledge, they will offer higher wages to avoid high worker turnover. Egger and Kreickemeier (2013) explain that more-productive firms which make higher profits will pay higher wages, regardless of ownership. In fact, they argue that if foreign and domestic companies have the same level of productivity there will be no foreign wage premium. They propose that wage premium is related to a company's global as opposed to national profits, and that it is not just that productive MNEs with FDI pay higher wages.

Fourth, Gorg et al. (2007) suggest that the foreign wage premium is a gradual process that relates to the accrual of skills by the workforce. This implies that the foreign wage premium is gained over time because workers become more valuable to the company after they have gone through substantial training and acquired job-specific experience. MNEs have been shown to be larger and more productive, to have better access to capital through their headquarters, and to have higher profits, which can also explain their higher investment in employees and greater expenditure on wages (Pearce, 2018).

These four underlying factors are all expected to positively affect average wages (Earle et al., 2012). Whether foreign ownership will have a positive or negative impact on individual workers' wages depends on human capital quality and on the presence of domestic MNE competition, which is stronger in developed than in developing countries. The most productive companies in developed countries can bear the costs of the liability of being foreign and are prone to become MNEs. Since domestic companies in developing countries are less productive than MNEs because they are further away from the technological frontier, the foreign wage premium can be seen as inevitable, to a greater or lesser extent. However, MNEs face greater competition in developed countries where domestic companies are closer to the technological frontier, so the foreign wage premium may be lower in developed countries. The technological gap between MNEs' home and host country plays a role in determining the size of the foreign wage premium. Consequently, it is no surprise that research in this field has shown that the impact of FDI on wages is, on average, neutral in developed countries and positive in developing countries (Javorcik, 2014).

Empirical literature

The empirical work on the foreign wage premium falls into three main categories. The first focuses on changes in company ownership while the workers remain in place, investigating whether changes in company ownership influence wages (Hijzen et al., 2013; Heyman et al., 2007). The second compares average wages in MNEs and non-MNEs (Heyman, 2007). The third focuses on the impact of company ownership on wage change by observing worker movement between companies and considering individual worker characteristics (Martins, 2011), and is the most closely related to this study. Although in explaining these phenomena we take the same broad approach as Martins, the methods used in

this research diverge from that study due to differences in data structure. Martins (2011) had data on domestic-to-domestic and foreign-to-foreign company change as well as firm or worker size and a measure of workers' experience, which is not available in the dataset from the Serbian Social Register.

Analysing the change of company ownership via foreign takeover in Sweden, Heyman et al. (2007) find that foreign ownership causes an 11% increase in wages at the company level, even after controlling for industry characteristics. However, company-level analysis is not perfect. As the authors acknowledge, foreign acquisitors tend to target domestic companies that already pay above-average wages. Because of this, Heyman et al. (2007) compare the wage difference between foreign-owned and domestic MNEs. They conclude that wage differences at the worker level are not driven by foreign ownership and that they exist between MNEs and non-MNEs rather than between domestic and foreign companies. This means that a large part of the discrepancy in wages between foreign and domestically owned companies is explained by their level of multinationality. The foreign ownership wage premium was then only 2% and the difference between foreign and Swedish-owned MNEs was almost zero.

Heyman (2007) shows that at the worker level, foreign companies pay about 4% higher wages than domestic companies, but when individual characteristics such as experience and education are accounted for the premium drops to about 2.5%. However, it has been argued that the majority of studies analyse the causal effect of change in ownership by observing domestic companies that have been acquired by MNEs and not by observing the employee changes that come with acquisition: the dismissal of unskilled workers and hiring of skilled workers, which can bias the foreign wage premium upwards (Hijzen et al., 2013).

In a comparative cross-country analysis of the UK, Germany, Portugal, Brazil, and Indonesia, Hijzen et al. (2013) study the impact of company ownership on wage changes in developed and developing countries separately, at both the company and the worker level. They confirm that in less-developed countries there is a greater difference between foreign and domestic wages. At the firm level the acquisition of a domestic company by a foreign company would lead to a 2% wage increase in Germany and a 21% wage increase in Indonesia. However, both Hijzen et al. (2013) and Heyman (2007) confirm that the estimated foreign wage

premium is significantly reduced at the worker level when individual characteristics are controlled for. Even after controlling for these factors, foreign ownership maintains its significant effect on movers' wages. Hijzen et al. (2013) attribute the foreign wage premium to the higher quality jobs provided by foreign companies.

Martins' analysis (2011), based on Portuguese data, is most closely related to the analysis in this paper. He finds that workers moving from domestic to foreign companies experience an average wage change of 18%, while those moving in the opposite direction experience an average wage change of –8.4%. The study focuses specifically on labour mobility (workers changing between foreign and domestic companies) rather than companies changing ownership. The same approach is used in this paper, as observing workers makes it possible to disentangle the different abilities of workers and the company's wages. However, workers may work for low-paying domestic companies, in which case a change from domestic to foreign company would most likely provide biased results. This possibility is controlled for by taking firm-specific characteristics into account. Moreover, the phenomenon of wage growth is not only about wage change at the moment of a job change, but also about wage growth in the long run. Martins (2011) also finds that workers who move from a domestic to a foreign company experience a higher wage growth (of about 4%) than those who move from a foreign to a domestic company (about 2%). Therefore, changing jobs from a domestic to a foreign company is expected to have a positive impact not only on wage change but also on wage growth. Similar results have been found for Germany and Norway, where moving from a domestic to a foreign company leads to a wage increase (Balsvik, 2011; Andrews et al., 2010).

Overall, there seems to be a consensus in the literature that a wage premium is associated with foreign ownership of a firm, mostly evident in developing countries. Therefore, based on the literature discussed thus far on the foreign wage premium, the first hypothesis is proposed:

H1: Changing from a domestic- to a foreign-owned employer is associated with increased wages.

2.3 Heterogeneous effects: the role of workers' education

Following Jovanovic (1979), one needs to allow for the fact that workers are heterogeneous, which means that they differ in productivity as in knowledge and skills. Therefore, wage growth should be related to workers' different abilities. At the same time, labour market information asymmetry results in uncertainty about labour productivity before employment. Consequently, as proposed by Campbell (2001:4), "the starting wage offered by companies is based on the expected value of productivity given the information available at the time the job commences. When starting the new job, there may initially be great uncertainty over actual productivity which implies that as new information arrives, future earnings may rise considerably above or below the starting wage".

Managing workers' wages is an important aspect of the human resource management that is responsible for the success of MNEs. An important aspect of MNEs' success is the ownership advantage that derives from knowledge; i.e., the firm's human capital. Therefore, human resource management is a core element of the advantage of MNE ownership. A company's human capital and financial performance are complementary (Narula & Verbeke, 2015). More-educated workers generate more knowledge and stronger ownership advantages for foreign companies; thus they are offered higher wages. This is one of the reasons for rewarding valuable workers, and the way MNEs appreciate employees has long-term consequences for the company's organisational strategy (Andersson et al., 2019).

Whether wages rise or fall after changing jobs depends on many measurable factors such as worker's education and experience with previous employers, and also on factors that cannot be measured precisely, such as a worker's skill match with the job requirements (Widerstedt, 1998). Observing wages and United States-based MNE activity in Mexico in the 20th century, Feenstra and Hanson (1997) find that FDI was positively correlated with the demand for highly skilled labour. The study observes the activity of regional manufacturing facilities, since a big share of employment was generated by the outsourcing of US MNEs. The regions with the biggest concentration of FDI were also the regions with the biggest increase in the share of skilled labour in total wages. According to Heyman et al. (2007), those who change companies also have work experience with the previous employer that might be valuable. Therefore, wage changes are driven by worker heterogeneity, or the 'heterogeneity effect'.

According to Gorg et al. (2007), by controlling for education and other worker characteristics the coefficient of foreign ownership is reduced but is positive and highly significant. A 1% increase in foreign ownership leads to an increase in the hourly wage of about 0.45%. Batra and Tan (2002) support this argument and also find that there are high productivity gains from MNE training. Not only is the training provided valuable for employees' work but it also provides them with personal satisfaction and the sense of being valued. Employee training results in productivity gains of up to 75% in Indonesia and Nicaragua and up to 45% in Mexico and Malaysia (Batra and Tan, 2002). This relationship between training and productivity gains has been confirmed in multiple studies (Ben Jamaa Cherif, 2021; Chhetri et al., 2018; Moussaid et al., 2020).

Research by Poole (2013), although predominantly focusing on the spillovers from MNEs to domestic companies through workers who move from one company to another carrying knowledge, social capital, or management style, also suggests that more-educated MNE employees benefit more from positive wage change than less-educated employees when they move to domestic firms. They argue that higher-skilled former MNE workers are better able to convey information and technology to domestic company workers and that higher-skilled domestic company workers are better able to absorb new technology through interactions with former MNE workers. The largest spillovers occur when former MNE workers have greater educational attainment and experience than domestic company workers (Poole, 2013)

Abbott and Beach (1994) find evidence that the education of workers matters for a wage increase when changing jobs. Those with university degrees experience higher wage change when changing jobs than workers with lower educational attainment. Their argument is that more-educated workers have greater ability to absorb knowledge in the workplace and use it productively in the new workplace. Therefore, the level of education moderates the relationship between job mobility and wages in addition to accumulated firm-specific human capital, and is also an important determinant of job change itself. Sousa-Poza and Henneberger (2004) find that more years of schooling lead to a higher propensity to change jobs because more-educated workers overcome the transition between jobs and adapt to a new environment more easily (Bowlus and Neuman, 2006). Therefore, the objective expectation is that education plays a critical role in wage growth (Mincer, 2012).

Based on the discussed literature about the role of individual worker characteristics and education in wage change, the second hypothesis is proposed:

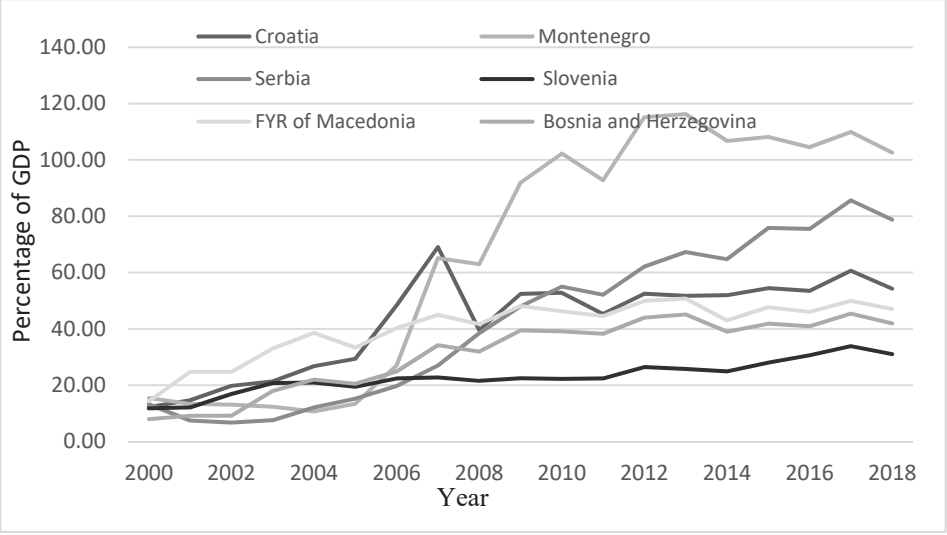
H2: More-educated workers benefit more than less-educated workers from changing from a domestic to a foreign-owned company.

3. METHODOLOGY AND DATA

3.1 Country and data characteristics

This empirical analysis of labour mobility and wage change in Serbia is based on employer–employee data for the period 2000 to 2014 from the Serbian Social Register. This is a particularly interesting period for this kind of analysis, since inward FDI in Serbia was negligible before 2000 and grew at a high rate after that due to political changes and liberalisation policies that favoured foreign capital, as Serbia developed the characteristics of a transition economy (Knežević et al., 2020; Petrovic et al., 2017). As presented in Figure 1, the flow of inward FDI to Serbia was among the highest in the region. With the exception of early post-NATO-conflict years when investment stagnated, there was a sharp increase in inward FDI to Serbia. After 2009, of the former Yugoslav Republic countries only Montenegro had higher inward FDI.

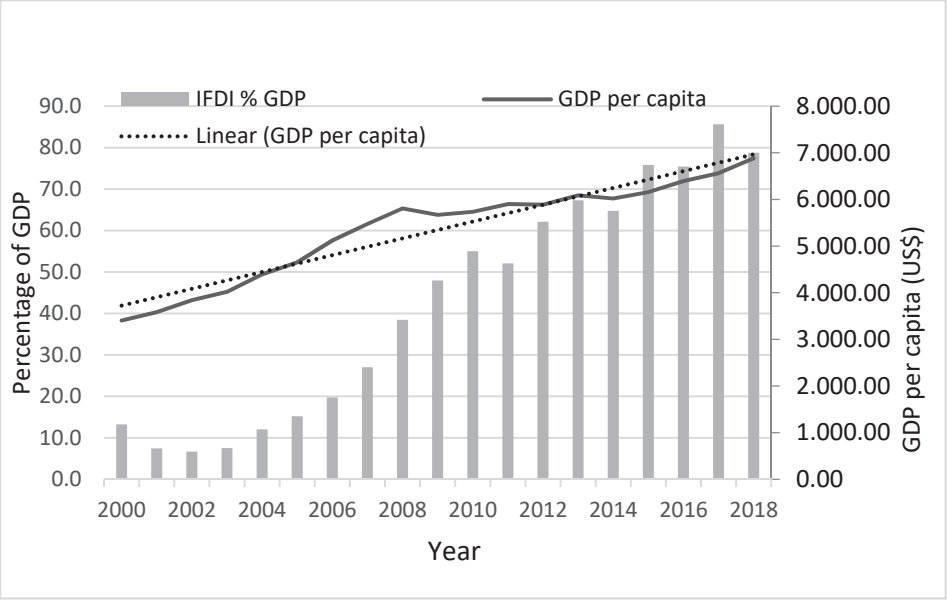
Figure 1: Inward FDI stock as % of GDP in ex-YU countries



Source: Author's illustration based on UNCTAD data

As shown in Figure 2, inward FDI stock to Serbia as a percentage of GDP started at about 13% in 2000 and reached almost 80% in 2018. However, average GDP per capita fluctuated slowly between around US$4,000 in 2000 and US$6,800 in 2018 – only a 33% increase over 16 years, giving an average annual growth of around 2%.

Figure 2: Serbia: Inward FDI stock, % of GDP and GDP per capita



Source: Author's illustration based on UNCTAD data

Serbia is an interesting case, as like many other former socialist countries its economic policies were based on FDI-driven economic growth. After the fall of Milošević in 2000, like other countries in the region, Serbia based its FDI policies on large subsidy packages. Conditions set by the Development Agency of Serbia[1] included the provision of urban construction sites, corporate tax exemptions, and a subsidy of up to EUR 10,000 for every new job.

This study focuses on employer–employee data which tracks workers over the period 2000 to 2014. This covers the history of inward FDI in Serbia, since MNEs

[1] http://ras.gov.rs/en/invest-in-serbia/why-serbia/financial-benefits-and-incentives

started investing in Serbia via FDI in 2000. The dataset allows observation of employees' gross annual wage and education level and employer ownership (foreign vs. domestic).[2] Observing the nominal wage would not have changed the pattern, so we did not need to deduct taxes from the gross wage provided by the Serbian Social Register. Entities are considered foreign if 10% or more is held by a foreign national, in line with International Monetary Fund criteria.

This paper does not analyse suggestions, for example, by Heyman (2007), that 'multinationality' matters and that the main wage differences are not primarily the result of differences between domestic and foreign-owned MNEs. This is not only because of data constraints but also because Serbia, as a transition country, does not have many privately owned domestic MNEs, a fact that is reflected in Serbia's extremely low outward FDI flows (UNCTAD, 2019).

The dataset consists of 1,500 individuals randomly chosen from people born between 1965 and 1975. No individual retired within the observed period. In 2014 Serbia had 3.1 million people active in the labour market, 2.5 million employed, 700,000 of them in the public sector, and an unemployment rate of 19%.

The sample excludes employees in the public sector and in state-owned enterprises. In the observed period the sampled individuals worked in around 3,000 private companies. None of the companies changed ownership via acquisition, so the focus is on workers changing company rather than companies changing ownership. The data allows identifying changes in ownership from foreign to domestic and from domestic to foreign, and investigating whether such changes led to changes in wages.

However, due to data limitations it is not possible to identify workers who changed jobs within the same type of company ownership, i.e., from domestic to domestic or foreign to foreign. Thus, the group of workers that do not change jobs could actually contain some that do: workers moving between domestic companies or between foreign-owned companies cannot be distinguished from workers staying in the same company. This is indeed a limitation. However, we do know that a substantial proportion of job-changers are accounted for. Based on current results, if all job-changers were considered, including those who

[2] See Appendix 2.1 for variable explanation and descriptive statistics

moved between companies with the same type of ownership, the wage difference between those who changed jobs and those who did not would potentially be even higher, although this is beyond the scope of this study.

The data allows for the measurement of not only the difference in wage growth between workers moving from domestic to foreign companies and vice versa, but also for comparison of the wages of movers from foreign to domestic companies with the wages of those who worked only in domestic firms.

The available data does not show the hourly/daily wage, the overall number of days an individual worked during a year, or the days a worker was employed in domestic and foreign companies. This is an issue, because it means that wages in the year workers changed jobs cannot be compared, as they received unspecified wages from both foreign-owned and domestic firms.

This limitation was dealt with by considering the growth in wages from one year before changing jobs to one year after changing jobs, thus excluding the year the workers change employer. Robustness checks were made using longer periods before and after changing jobs. The data cleaning procedure involved removing observations with extreme values for wage change (over 3000%). Some individuals were missing wage data for the year after the change. Therefore, the final sample comprised 984 individuals.

Every individual's wage was observed in relation to the company employing them. The wage change *Y* is computed as the percentage change between the wage one year before and one year after changing company. In this paper we are interested in exploring the 'foreign-wage premium' phenomenon, not the domestic-wage premium that would occur when changing from a foreign to a domestic company. Therefore, the variable *DF* – domestic to foreign – is observed as the independent variable rather than *FD* – foreign to domestic. The dummy variable *DF* takes value 1 in a year of change from a domestic to a foreign company and value 0 in a year of no change.

On the other hand, the variable change *N*, the relationship between changing jobs and wages, is constructed regardless of whether it is from foreign to domestic company or vice versa. The variable *N* takes the value 1 if the worker changed employer and 0 otherwise.

The change in wages is also computed for a group of individuals who did not change type of company. They form a control group, which means they only worked for foreign or domestic companies but their wage change is observed with respect to workers who did change type of employer. Thus, two additional variables are created for foreign company workers only (*F*) and domestic company workers only (*D*). Variable *F* takes the value 1 if individuals worked for a foreign company only and 0 otherwise, and *D* takes the value 1 if individuals worked for a domestic company only and 0 otherwise. Their wage change is observed on a year-by-year basis.

The education *E* of workers is based on five education levels. The education variable takes value 1 – primary school, 2 – secondary school, 3 – upper secondary, 4 –Bachelor's degree, 5 – Master's degree. The median education of the worker is computed for the whole observed period and used as a time-invariant variable.

3.2 Econometric Analysis

This section presents the results of the econometric analysis. The role of this analysis is to estimate how the wage dynamic differs between workers who moved from domestic to foreign-owned firms or from foreign-owned to domestic firms, and those who remained in either domestic or foreign-owned firms.

The purpose of Equation (1) is to investigate whether changing jobs means a change in wages:

$$Y_{it} = \alpha + \beta_1 N_{it} + \delta_1 E_i + \varepsilon_{it} \quad (1)$$

where *Y* is the wage change for worker *i* at the time *t*, *A* is constant, and *N* takes value 1 if the worker changed companies, and 0 otherwise. *E* is a measure of workers' education level, constructed as an ordinal variable taking values from 1 to 4, where 1 = primary school, 2 = secondary school, 3 = undergraduate degree, 4 = Master's degree. ε_{it} is the error term.

The purpose of Equation (2) is to investigate whether wage change at time *t* (between t–1 and t+1) is driven by company ownership:

$$Y_{it} = \alpha + \beta_1 DF_{it} + \delta_1 E_i + \varepsilon_{it} \quad (2)$$

In Equation (2) *DF* takes value 1 if the worker moved from domestic to foreign-owned company within a year, and 0 otherwise.

Equation (3) further distinguishes the *N*-type of workers who remain in domestic firms *D* or in foreign firms *F:*

$$Y_{it} = \alpha + \beta_1 DF_{it} + \beta_2 F_{it} + \beta_3 D_{it} + \delta_1 E_i + \varepsilon_{it} \quad (3)$$

where *D* takes value 1 if the worker remained in domestic companies, and 0 otherwise, and *F* takes value 1 if the worker remained in foreign companies, and 0 otherwise.

In order to test whether the level of education moderates the effect of worker mobility on wage change, we estimate the following Equation (4):

$$Y_{it} = \alpha + \beta_1 N_{it} + \beta_2 DF_{it} + \delta_1 E_i + \beta_3 DF_{it} * E_i + \varepsilon_{it} \quad (4)$$

The coefficient β_3, associated with the interaction between education and change from domestic to foreign company dummy (*DF*E*) captures whether more-educated workers benefit more from moving from domestic to foreign company.

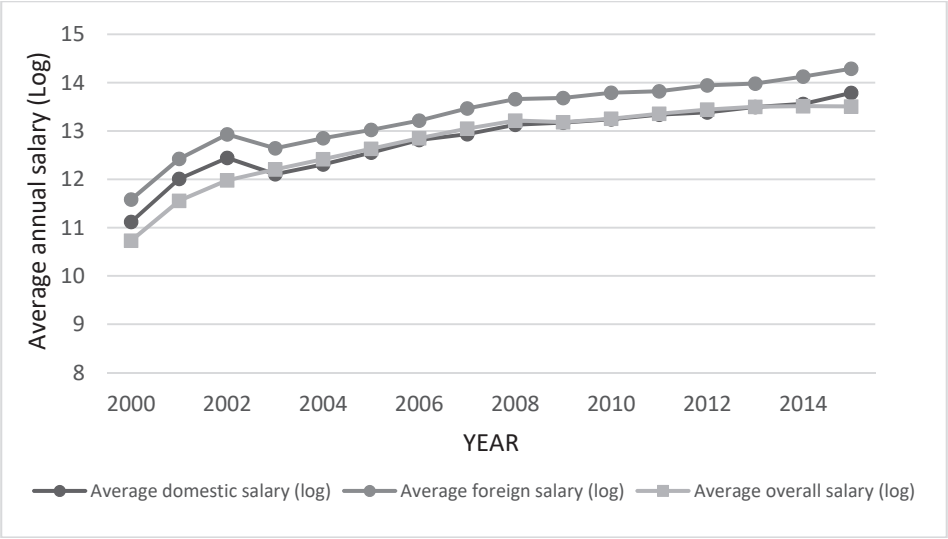
Regressions are run where the main explanatory variable *FD* is substituted with *DF=1–FD*. This allows for the presentation of results in a different form, to show whether workers moving from foreign to domestic firms enjoy a wage discount, as opposed to those staying in foreign or domestic firms.

4. RESULTS

4.1 Descriptive statistics

The descriptive data overview shows that, on average, MNEs pay higher wages than domestic companies. As presented in Figure 3, over the period 2000 to 2015 MNE wages were consistently higher than those paid by domestic companies, with national average wages fluctuating somewhere between the two.

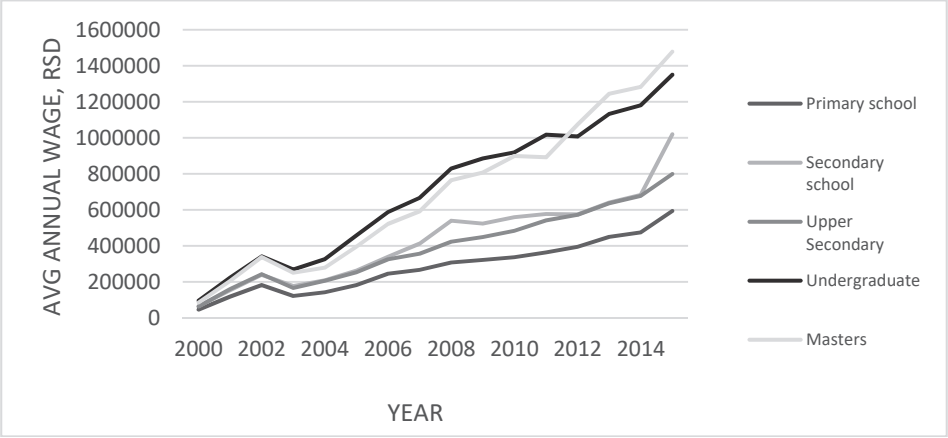
Figure 3: Salary average (log) – domestic vs. foreign



Source: Author's illustration based on Serbian Social Register data

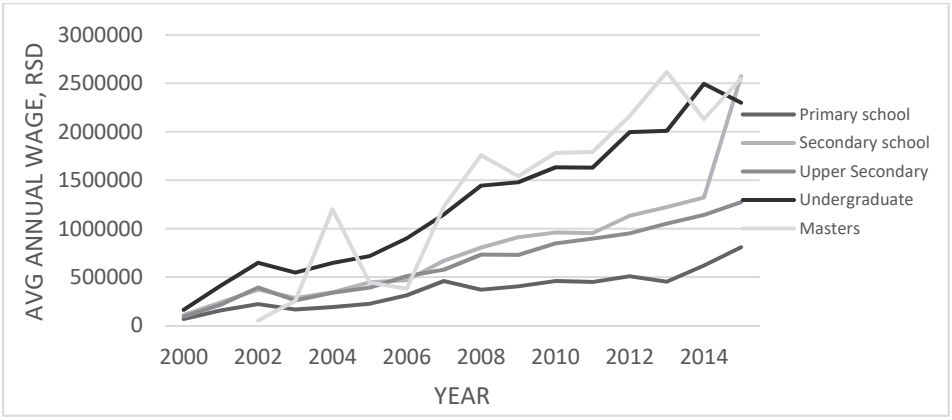
However, if observed individually (Figure 4 and Figure 5), in both types of company more-educated workers receive higher wages than less-educated workers.

Figure 4: Average wages in domestic companies, by education level



Source: Author's illustration based on Serbian Social Register data

Figure 5: Average salaries in foreign companies, by education level



Source: Author's illustration based on Serbian Social Register data

Over the observed period (2000 to 2015) the vast majority of workers that changed employer (about 83%) did it only once during their career. Furthermore, local firms were observed to experience an influx of former MNE employees during this period: in the observed sample, 63.4% of workers changed from a foreign to a domestic company. On the other hand, 22% changed from a domestic to a foreign company. This is based on the observation of employee movement between different ownership categories within the given sample, rather than on an investigation of the causes for this particular trend where the majority of workers move from foreign to domestic companies. The observed sample shows such a trend but the subject of this investigation is whether those movements led to a change in wages.

4.2 Regression Analysis

The regression output shows that the model has high explanatory power. About 18% of the variation in wages is explained by changes from domestic to foreign companies. According to Table 1 column (1), changing jobs already has a positive impact on wages, regardless of the direction: the wage growth of those who change jobs is about 80% higher than for those that do not change. However, as shown in column (2), that impact is even greater when the change is from domestic to foreign company. On average, the change in the wages of workers moving from domestic to foreign firms is over 90% higher than the change in the wages of workers that either do not change employer or move from foreign to

domestic companies. The emphasis here is not on wages doubling but on them having higher growth. For example, if the wages of those who change from foreign to domestic companies increases by 2%, the wages of those who change from domestic to foreign companies will increase by 4.2%. Hence, HP 1 is confirmed.

Table 1: The effect of changing jobs on wages

	(1)	(2)	(3)	(4)
Variable			**Y**	
N	83.61***			61.42***
	(7.591)			(13.11)
DF		92.61***	32.30**	–13.99
		(9.081)	(15.91)	(34.19)
F			–53.60***	
			(14.31)	
D			–62.21***	
			(13.32)	
E	6.461***	6.401***	6.605***	6.517***
	(2.312)	(2.314)	(2.315)	(2.241)
DF*E				21.28*
				(13.98)
Year dummies	yes	yes	yes	yes
Constant	–85.24***	–85.24***	–24.20	–99.70***
	(14.86)	(14.87)	(19.82)	(15.66)
Observations	14,756	14,756	14,756	14,756
R-squared	0.187	0.186	0.188	0.188

Standard errors in parentheses

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

When controlling for the wage growth of non-changers in column (3), it is noticeable that changing jobs from a domestic to a foreign company leads to a positive wage change compared to those who stay in foreign or domestic

companies only. Those workers that change from domestic to foreign companies have a 32% higher wage change compared to those who change from a foreign to a domestic company.

Finally, as shown in column (4), education has a direct effect that is unrelated to changing jobs and change of ownership. The change in wages is greater for workers with higher education, independent of whether they change jobs. This also means that the wages of more-educated workers and less-educated workers diverge. The results in column (4) show that education has a positive and significant moderating effect. Workers with the lowest education level experience the least positive effect of job change, while workers with the highest educational attainment experience higher wage increases which confirms HP2 This supports the view that MNEs reward higher-educated workers more, as they are key to maintaining and developing the ownership advantage of MNEs.

Interestingly, the change from domestic to foreign company is completely insignificant in column (4), as all the effect is captured by the interaction of that change with education. In other words, for workers with education below secondary education (E=2), changing jobs has no effect.

4.3 Robustness checks

The wage change in Table 1 was observed one year before and one year after changing company. In order to test the validity of this observation, the robustness test estimates the impact of changing company on an average wage change, two years before and two years after (Table 2). The results remain qualitatively the same. A change from domestic to foreign company causes a significant change in wages, and the more educated the worker the higher the wage change. However, in this case the direction of change (domestic to foreign) matters and not the change itself. If the same analysis is conducted on the basis of a wage change that is computed as an average change three years before and three years after changing companies, the results remain very similar. Although this reduces the number of observations, it confirms the robustness of the analysis.

Table 2: The effect of changing jobs on wages (2-year window)

	(1)	(2)	(3)	(4)
Variable	**Y** (2-year window)			
N	77.35***			29.13***
	(4.960)			(8.706)
DF		98.41***	69.93***	–28.95
		(5.901)	(10.40)	(22.83)
F			–28.60***	
			(9.471)	
D			–29.05***	
			(8.737)	
E	7.439***	7.439***	7.475***	7.976***
	(1.287)	(1.287)	(1.289)	(1.610)
DF*E				45.07***
				(9.319)
Year dummies	yes	yes	yes	yes
Constant	8.984	8.984	37.98***	–8.724
	(8.974)	(8.960)	(12.51)	(9.620)
Observations	10,363	10,363	10,363	10,363
R-squared	0.100	0.102	0.103	0.108

Standard errors in parentheses

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

5. CONCLUSIONS

This paper contributes to the empirical literature on wage change and foreign ownership by studying the case of Serbia, using rich employer–employee data. Previous research on labour mobility and wage change shows that job change is not always related to increased wages. There are many additional factors that influence wage change, including demographic characteristics. Some previous studies have suggested that change within rather than across occupation is important. However, the literature also shows that company ownership and

worker characteristics are independent drivers of wage change. The foreign wage premium has been mostly observed in developing countries because domestic competitors lack the knowledge and technology to compete with MNEs (Coniglio et al., 2015; Hijzen et al., 2013; van der Straaten et al., 2020).

The findings of this study are in line with other studies on wage change and FDI in developing countries: MNEs in Serbia not only pay higher wages on average, but they also pay higher wages to similar workers, so the change from a domestic to a foreign company alone leads to a wage change. Although the lessons from this study are based on Serbian data, the results are in line with those obtained for other countries, so the findings apply more generally.

At the same time, data limitation makes it impossible to identify whether workers who change jobs within the same type of company ownership (domestic to domestic or foreign to foreign) also experience a salary change. Furthermore, other worker and company characteristics (such as worker experience with a previous employer or company industry) are not observed, so strong causality as such cannot be claimed. As Gorg et al. (2007) suggest, even when observing different worker and company characteristics and accounting for training provision, it is hard to isolate the specific causes of the foreign wage premium since factors like experience, social capital, and learning by doing are hard to measure.

These results do not simply suggest that MNEs pay higher wages on average because they simply employ a greater proportion of skilled workers, as some previous studies have proposed. More-educated workers who change from domestic to foreign companies benefit more from such a change than less-educated workers. This is in line with Heyman's (2007) finding that a higher education level has a positive impact on average wage change. Kidd (1991) also confirms the positive relationship between years of schooling and the probability of changing companies. More-educated workers who move from a domestic to a foreign company experience a 21% higher wage change.

It is important to highlight that job change may be the result of dismissal or voluntary change. However, the dataset used in this study does not provide information on dismissals. It is possible that the foreign wage premium is driven by the fact that the foreign company dismisses workers who then have to accept

a lower salary in a domestic company. However, other foreign employers are available to dismissed workers.

This paper also shows that the wage increases are higher for workers with higher education, independent of changing jobs. On top of this, the wage gain from moving to foreign MNEs is concentrated in the more-educated workers. This means that the wages of more-educated workers and less-educated workers diverge.

Although there is a strong correlation between increased wages and changing companies, it is not known if this is caused by asymmetric labour market information, higher productivity within MNEs, or higher investment in training. The argument that training is a driver of higher wages in MNEs has been made in previous studies but is not a variable controlled for in this paper.

The findings have important policy implications. First, attracting MNEs can increase the wages of Serbian workers, thus increasing welfare and consumption and boosting Serbian economic growth. However, since these gains will go disproportionally to more-educated workers, the resulting wage inequality and the potential tensions that this could create need to be considered (Alili and Adnett, 2018). Figini and Gorg (2011) show that wage inequality in developing countries increases with inward FDI stock (as a percentage of GDP). Hale and Xu (2016) suggest that this is mostly due to FDI bringing more sophisticated technologies and managerial practices to secondary industries, which demands more-educated workers.

Consequently, a higher demand for more-educated labour leads to higher wages for this group of workers, creating a gap between the more- and less-educated. However, this effect diminishes as countries approach the technological frontier. By measuring total wage inequality,[3] Figini and Gorg (2011) find that while FDI, on average, increases wages in host countries and makes some workers better off in absolute terms, this undermines the balance in wages between skilled and unskilled workers (Figini and Gorg, 2011).

[3] By means of Gini and Theil inequality indices

Second, the evidence is consistent with the idea that attracting MNEs can leverage investment in education and support knowledge transfer (Park et al., 2021; Zidan, 2001). The larger the share of workers with a high level of education, the larger the benefits from attracting MNEs. However, while wage increase is good news for Serbian workers, it could have an adverse effect on local Serbian companies, which are likely to face the prospect of their best workers moving to MNEs, or having to pay higher wages in order to retain their workers. Without a corresponding increase in productivity, this may severely harm the competitiveness of Serbian companies.

A more precise answer to the question of the impact of MNEs on labour could be provided by observing worker development within MNEs. In particular, the progress of employees in the corporate hierarchy would reveal more than just the relevance of their education to the company. The foreign wage premium might be accompanied by other benefits like training or by negative conditions such as blocked ability to progress to managerial positions. Employing local labour in managerial positions in foreign companies would indicate that MNEs are contributing to higher wages. Career progress and eventual pay rises over time would suggest that employees gain valuable experience, in addition to their education. Possible directions for future research include observing the role of current SME owners' experience gained in MNEs, and looking at the spinoff effect of the presence of MNEs in the labour market, rather than just the financial effect.

REFERENCES

Abbott, M.G., & Beach, C.M. (1994). Wage changes and job changes of Canadian women: evidence from the 1986–87 labour market activity survey. *Journal of Human Resources*, 29(2), 429–460.

Alili, M.Z., & Adnett, N. (2018). Did FDI increase wage inequality in transition economies? *International Journal of Social Economics*, *45*(9), 1283–1304.

Aitken, B., Harrison, A., & Lipsey, R. (1996). Wages and foreign ownership: A comparative study of Mexico, Venezuela, and the United States. *Journal of International Economics*, 40, 345–371.

Andrews, M., Bellmann, L., Schank, T., & Upward, R. (2010). The takeover and selection effects of foreign ownership in Germany: an analysis using linked worker–firm data. *Review of World Economics/WeltwirtschaftlichesArchiv,* 145(2), 293–317.

Ayentimi, D. T., Burgess, J., & Brown, K. (2018). HRM practices of MNEs and domestic firms in Ghana: divergence or convergence? *Personnel Review*, *47*(1), 2–21.

Batra, G., & Tan, H. (2002). Upgrading work force skills to create high-performing firms. In I. Nabi, & M. Luthria (Eds), *Building Competitive Firms: Incentives and Capabilities.* Washington DC: World Bank Publications, pp.118–134.

Balsvik, R. (2011). Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing. *The Review of Economics and Statistics*, *93*(1), 285–297.

Bartel, A.P., & Borjas, G.J. (1981). Wage growth and job turnover: An empirical analysis. In S. Rosen (Ed.), *Studies in Labor Markets* (pp. 65–90). Chicago: University of Chicago Press.

Bernard, A.B., & Sjoholm, F. (2003). *Foreign owners and plant survival* (Working Paper 10039). National Bureau of Economic Research. https://www.nber.org/papers/w10039

Blonigen, B.A., & Slaughter, M.J. (2001). Foreign-affiliate activity and US skill upgrading. *The Review of Economics and Statistics*, *83*(2), 362–376.

Blomstrom, M. (2002). The economics of international investment incentives. *International Investment Perspectives*, 1, 165–183.

Bossler, M., & Broszeit, S. (2017). Do minimum wages increase job satisfaction? Micro-data evidence from the new German minimum wage. *Labour*, *31*(4), 480–493.

Ben Jamaa Cherif, N. (2021). The impact of training on productivity and wages in Tunisia. *International Journal of Manpower.* https://doi.org/10.1108/IJM-12-2019-0527

Chhetri, P., Gekara, V., Manzoni, A., & Montague, A. (2018). Productivity benefits of employer-sponsored training: A study of the Australia transport and logistics industry. *Education & Training*, *60*(9), 1009–1025.

Coniglio, Nicola D., Francesco Prota, and Adnan Seric (2015). Foreign direct investment, employment and wages in Sub-Saharan Africa. *Journal of International Development* 27(7), 1243–1266.

Cruz, M., Nayyar, G., Toews, G., & Vezina, P. L. (2018). *FDI and the skill premium: Evidence from emerging economies* (Policy Research Working Paper 8613). Washington DC: The World Bank.

Davis, C. A., & Poole, J. P. (2020). Can multinational firms promote gender equality? The role of labour mobility. *Transnational Corporations Journal*, 27(3), 157-181.

Dewit, G., Görg, H., & Temouri, Y. (2019). Employment protection and firm relocation: Theory and evidence. *Economica*, *86*(344), 663–688.

Dobbelaere, S., & Kiyota, K. (2018). Labor market imperfections, markups and productivity in multinationals and exporters. *Labour Economics*, *53*, 198–212.

Doh, J. (2019). MNEs, FDI, inequality and growth. *Multinational Business Review*, *27*(3), 217–220.

Bunzel, H., Christensen, B., J., Neumann, G. R., & Robin, J.M. (Eds.) (2006). *Structural Models of Wage and Employment Dynamics*. Amsterdam: Elsevier.

Campbell, D. (2001). *Estimating the wage effects of job mobility in Britain* (Studies in Economics 0117). Department of Economics, University of Kent.

Dunning J. (1998). *The changing geography of foreign direct investment: Explanations and implications*. In N. Kumar (Ed.), Globalization, Foreign Direct Investment and Technology Transfers: Impacts on and Prospects for Developing Countries. London: Routledge.

Earle, J., Telegdy, Á., & Antal, G. (2012). *FDI and wages: Evidence from firm-level and linked employer-employee data in Hungary, 1986–2008* (IZA Discussion Paper no. 7095). Bonn: Institute of Labor Economics

Egger, H., & Kreickemeier, U. (2013). Why foreign ownership may be good for you. *International Economic Review*, *54*(2), 693–716.

Feenstra, R.C., & Hanson, G.H. (1997). Foreign direct investment and relative wages: Evidence from Mexico's maquiladoras. *Journal of International Economics*, *42*(3), 371–393.

Fosfuri, A., Motta, M., Rønde, T. (2001). Foreign direct investment and spillovers through workers' mobility. *Journal of International Economics*, 53 (1), 205–222.

Gereffi, G., Frederick, S., & Bamber, P. (2019). Diverse paths of upgrading in high-tech manufacturing: Costa Rica in the electronics and medical devices global value chains. *Transnational Corporations*, *26*(1), 1–30.

Girma, S., Gorg, H. (2007). Evaluating the foreign ownership wage premium using a difference-in-differences matching approach. *Journal of International Economics* 72 (1), 97–112.

Gorg, H., Strobl, E. (2003). Footloose multinationals. *The Manchester School,* 71(1), 1–19.

Gorg, H., Strobl, E., & Walsh, F. (2007). Why do foreign-owned firms pay more? The role of on-the-job training. *Review of World Economics*, *143*(3), 464–482.

Hahn, C.H., Hayakawa, K., & Ito, T. (2013) Do local national managers improve multinationals' performance? Evidence from Korean firm-level data. In K. Hayakawa (Ed.), *Deepening of Corporate Global Activities in East Asia*, Bangkok: IDE-JETRO.

Heinrichs, K., Angerer, P., Li, J., Loerbroks, A., Weigl, M., & Müller, A. (2020). Changes in the association between job decision latitude and work engagement at different levels of work experience: A 10-year longitudinal study. *Work & Stress*, *34*(2), 111–126.

Heyman, F., Sjoholm, F., Gustavsson Tingvall, P. (2007). Is there really a foreign ownership wage premium? Evidence from matched employer–employee data. *Journal of International Economics* 73 (2), 355–376.

Hijzen, A., Martins, P.S., Schank, T., & Upward, R. (2013). Foreign-owned firms around the world: A comparative analysis of wages and employment at the micro-level. *European Economic Review*, *60*, 170–188.

Javorcik, B.S. (2014). Does FDI bring good jobs to host countries? *The World Bank Research Observer, 30*(1), 74–94.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, *87*(5, Part 1), 972–990.

Kidd, M.P. (1991). An econometric analysis of interfirm labour mobility. *Canadian Journal of Economics*, 517–535.

Knežević, M. N., Hadžić, O., Nedeljković, S., & Kennell, J. (2020). Tourism entrepreneurship and industrial restructuring: GLOBE national and organisational cultural dimensions. *Journal of the Geographical Institute Jovan Cvijić SASA*, *70*(1), 15–30.

Knežević, M.N., Petrović, M.D., Kovačić, S., Mijatov, M., Vuković, D.B., & Kennell, J. (2021). Acting the part: Emotional intelligence and job satisfaction as predictors of emotional labor in travel agencies. *Tourism and Hospitality Research*, *21*(2), 183–201.

Kupets, O. (2016). Education–job mismatch in Ukraine: Too many people with tertiary education or too many jobs for low-skilled? *Journal of Comparative Economics,* 44(1), 125–147.

Lisi, G. (2018). Job satisfaction, job match quality and labour supply decisions. *International Review of Economics*, *65*(4), 489–505.

Longhi, S., & Brynin, M. (2010). Occupational change in Britain and Germany. *Labour Economics*, *17*(4), 655–666.

Lipsey, R., & Sjöholm, F. (2004). Foreign Direct Investment, Education, and Wages in Indonesian Manufacturing. *Journal of Development Economics* 73(1), 415-422.

Martins, P.S. (2004). *Do foreign firms really pay higher wages? Evidence from different estimators* (IZA Discussion Paper 1388). Bonn: Institute of Labor Economics.

Martins, P.S. (2011). Paying more to hire the best? Foreign firms, wages, and worker mobility. *Economic Inquiry*, *49*(2), 349–363.

Masso, J., & Vahter, P. (2019). Knowledge transfer from multinationals through labour mobility: are there effects on productivity, product sophistication and exporting? *Emerging Markets Finance and Trade*, *55*(12), 2774–2795.

Moore, F. (2007). Work–life balance: contrasting managers and workers in an MNC. *Employee relations*, *29*(4), 385–399.

Moussaid, A., Tkiouat, M., & Hlyal, M. (2020). Review of literature on existing models about the impact of continuous training on business performance. *IOP Conference Series: Materials Science and Engineering*, 827(1), 1-6.

Mincer, J. (2012). Wage changes in job changes. In S. W. Polachek, & K. Tatsiramos (Eds.), *35th Anniversary Retrospective (Research in Labor Economics, Vol. 35)*, Bingley: Emerald Group Publishing Limited, pp. 467-493..

Miyamoto, K. (2003). *Human capital formation and foreign direct investment in developing countries* (Working Paper No. 211). Paris: OECD Development Center.

Narula, R. (1996). *Multinational Investment and Economic Structure: Globalisation and Competitiveness*. London: Routledge

Navaretti, G.B., Venables, A., & Barry, F. (2004). *Multinational Firms in the World Economy*. Princeton: Princeton University Press.

Pandit, B. L., & Siddharthan, N. S. (2020). MNEs, product differentiation, skills and employment: lessons from the Indian experience. In *High-Tech Industries, Employment and Global Competitiveness* (pp. 165–179). London: Routledge India.

Park, J.S., Chang, J.Y., & Lee, T. (2021). The impacts of inward knowledge transfer and absorptive capacity on the turnover of host country nationals in MNE subsidiaries: a multilevel modeling approach. *Journal of Knowledge Management*. https://doi.org/10.1108/JKM-03-2021-0182

Pearce, R. (2018). The Dynamics of Differentiation: The Resource Bases of Development and the Roles of MNE Subsidiaries. In *Contemporary Issues in International Business* (pp. 153–170). Cham: Palgrave Macmillan.

Petrović, M. D., Blešić, I., Vujko, A., & Gajić, T. (2017). The role of agritourism's impact on the local community in a transitional society: A report from Serbia. *Transylvanian Review of Administrative Sciences*, *13*(50), 146–163.

Poole, J.P. (2013). Knowledge transfers from multinational to domestic firms: Evidence from worker mobility. *Review of Economics and Statistics*, *95*(2), 393–406.

Pradhan, J.P. (2006). *Quality of foreign direct investment, knowledge spillovers and host country productivity: a framework of analysis. institute for studies in industrial development* (ISID Working Paper No. 2006/11). New Delhi: Institute for Studies in Industrial Development.

Sahu, S.K., & Goel, A. (2021). Wages and firm ownership: A study of the manufacturing sector of India. *Studies in Microeconomics*. https://doi.org/10.1177/23210222211051453

Sousa, N. (2001). Multinationals and technology transfer through labour training. In *CEPR Workshop on Labour Market Effects of European Foreign Investsments. Dublin*.

Sousa-Poza, A., & Henneberger, F. (2004). Analyzing job mobility with job turnover intentions: An international comparative study. *Journal of Economic Issues*, *38*(1), 113–137.

van der Straaten, K., Pisani, N., & Kolk, A. (2020). Unraveling the MNE wage premium. *Journal of International Business Studies*, *51*(9), 1355–1390.

Webbink, D., & Hartog, J. (2004). Can students predict starting salaries? Yes! *Economics of Education Review*, *23*(2), 103–113.

Widerstedt, B. (1998). *Moving or staying? Job mobility as a sorting process.* (Doctoral dissertation, Umeå Universitet).

Woodgate, R. (2021). *Profit-led in effect or in mere appearance? Estimating the Irish demand regime given the influence of multinational enterprises* (No. 154/2021). Berlin: Institute for International Political Economy.

Zidan, S.S. (2001). The role of HRD in economic development. *Human Resource Development Quarterly*, *12*(4), 437–443.

Received: July 18, 2021
Accepted: March 16, 2022

APPENDIX

2.1: Descriptive statistics and variable explanation

Variable	Label	Obs	Mean	Std. Dev.	Min	Max
Salary change	Y	14,756	57.08401	211.6524	–100	2979.677
Domestic to foreign company change	DF	14,756	0.0321903	0.1765112	0	1
Foreign company workers only	F	14,756	0.0879642	0.2832525	0	1
Domestic company workers only	D	14,756	0.8652074	0.3415135	0	1
Change of company	N	14,756	0.0468284	0.2112783	0	1
Education	E	14,756	2.144517	0.7101322	1	4
Education interaction with domestic to foreign company change	E*DF	14,756	0.0700393	0.4006065	0	4

t – year, j – company, i – worker

Correlation coefficients

Variable	Y	N	DF	F	D	E	E*DF
Y	1						
N	0.0969	1					
DF	0.0935	0.8228	1				
F	–0.001	–0.0688	–0.0566	1			
D	–0.0591	–0.5616	–0.4621	–0.7868	1		
E	0.0269	0.0049	0.008	–0.0511	0.0394	1	
E*DF	0.0952	0.7888	0.9587	–0.0543	–0.443	0.0534	1

ECONOMIC ANNALS, Volume LXVII, No. 232 / January – March 2022
UDC: 3.33 ISSN: 0013-3264

https://doi.org/10.2298/EKA2232081A

*Rochna Arora**
*Baljit Kaur*****

IS URBANISATION SANS INFRASTRUCTURE A MYTH? EVIDENCE FROM INDIA

ABSTRACT: *This study examines the determinants of urbanisation in Indian states with special emphasis on infrastructure and infrastructure investment, using data on 17 Indian states for 1991 to 2017. The fixed effects regression model shows that physical infrastructure is an important determinant in high income states, while social infrastructure is important in high-income and low-income states (where the magnitude is negative). Electricity consumption and teledensity positively affect urbanisation in high- and low-income states, while the infant mortality rate in high-income states and the enrolment ratio in low-income states affect urbanisation negatively. The supply-led inverted-U hypothesis of infrastructure-investment-led urbanisation is only disproven for middle-income states while applying strongly in all other cases, particularly low-income states. Hence, the impact of infrastructure on urbanisation differs across states not only by type of infrastructure but also by the state's income category.*

KEY WORDS: *infrastructure, urbanisation, India*

JEL CLASSIFICATION: O18, C38, G21, L91, Q40, O43

* Amritsar Law College, Amritsar, Punjab, India, email: rochna.arora17@gmail.com, (Corresponding Author)

** Punjab School of Economics, Guru Nanak Dev University, Amritsar, Punjab, India, email: baljit17_kaur@yahoo.co.in

No country has grown to middle income without industrialising and urbanising. None has grown to high income without vibrant cities. The rush to cities in developing countries seems chaotic, but it is necessary.

Angel et al. 2010

1. INTRODUCTION

Urbanisation is often regarded as a major driver of higher productivity and economic growth. Better employment opportunities, a better work environment, and a better lifestyle attract people to urban areas, contributing to the growth of urban centres and to urban incomes. Development of an ever-increasing number of cities drives economic growth through economies of scale in infrastructure, labour, and capital, amplifying the contribution of cities to overall economic growth. The popular Tiebout hypothesis (1956) posits that better provision of public facilities such as education and roads attract not only more residents but also more businesses, thereby accelerating city growth. The Indian experience of the cities Bangalore, Pune, and Hyderabad suggests that the accumulation of human capital caused the cities to grow by attracting firms to their skill pools (Black and Henderson 1999; Li and Cheng 2006; Bertinelli and Strobl 2007; Leitão 2013; Sekkat 2013; Liu et al. 2015; Quintana and Royuela 2014; Quintana 2017).

Urbanisation as a process also propels economic growth by transforming the economy from agriculture-driven to non-agricultural-based (Henderson 2003). Urbanisation per se may be less significant for economic growth because to some extent it depends on an existing enabling environment in the form of institutions (Turok et al. 2013). These institutions might take the form of infrastructure or governance, or both. Thus, public investment in infrastructure plays a crucial role in urbanisation. Zhang (2002) has shown that economic growth together with urban policy reforms, Foreign Direct Investment, and structural changes significantly contributed to rising urbanisation in China, especially post-reforms. Chakravorty (2007) also conjectured that the increased urbanisation of India's coastal cities was the result of a high level of Foreign Direct Investment (FDI).

There are three main channels of the phenomenon of growing urbanisation in India (Table 1). The first and most important is the increasing size of the population. Population increase can be defined as the difference between the

crude birth rate and crude death rate: if the birth rate is higher than the death rate the population increases, while if the death rate exceeds the birth rate the population shrinks.

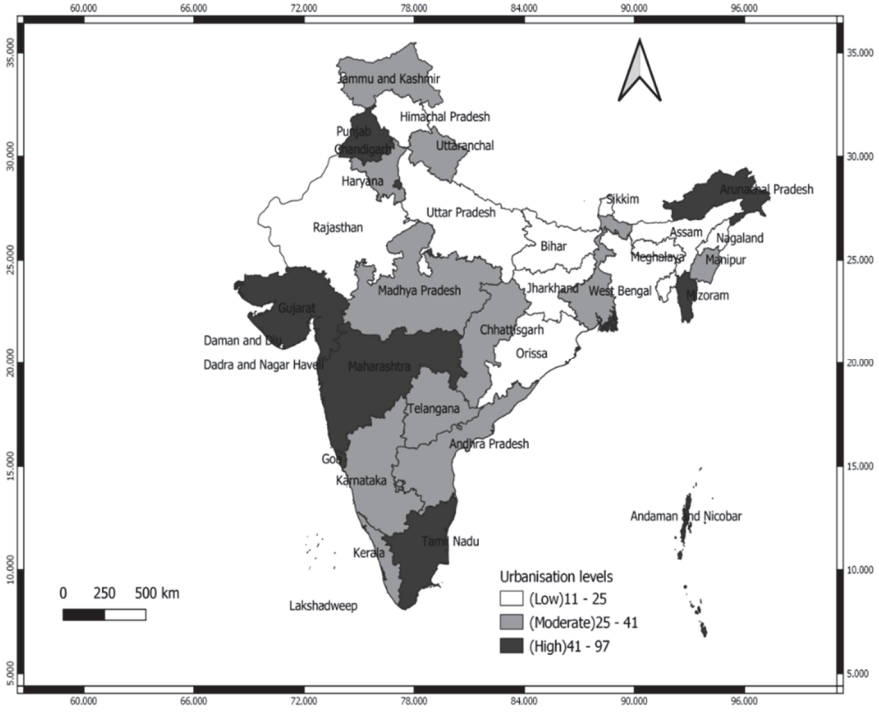
Table 1: Share of Urban Population in India

	1901	1951	1991	2001	2011
Total population (in millions)	238	361	846	1,028	1,211
Urban share (in %)	10.84	17.30	25.72	27.71	31.15

Source: Colmer 2017.

The second channel is the emergence of new urban areas, known in India as Urban Agglomerations. The third channel is demographic change: the migration of the rural populace to urban areas. The pace of urbanisation in India is not uniform across the country. Figure 1 below shows the diverse spread of urbanisation across Indian states.

Figure 1: Urbanisation in the Indian sub-continent 2019, in percentages



The urbanisation levels in the Indian states vary between 65.2% in Goa to as low as 10.6% in Bihar. The highest urbanisation rate is in Delhi (96.7%) and the lowest in Daman and Diu (20.9%).

These different rates of urbanisation bring into question the role played by infrastructure in the urbanisation process in Indian states. If infrastructure is an influence, is it uniform across states, or do states differ according to income? A number of papers have examined the determinants of urbanisation at the national level as will be discussed in the next section, but none has addressed this issue at the sub-national level, and none has addressed determinants at the sub-national level after segregating the states according to income. The present research not only discusses the determinants of urbanisation at the national level but also examines the asymmetrical importance of infrastructure across states according to different income brackets. Also, previous research has only used panel regression at one time point, whereas this paper uses urbanisation variables from sources other than the census.

The following section 2 reviews the relevant literature. Section 3 presents the database and methodology. The results are given in section 4, and the conclusions and policy implications in section 5.

2. LITERATURE REVIEW

While the earlier theories of Todaro (1969) and Harris and Todaro (1970) posited the rural–urban wage differential as a major factor behind urbanisation, Krugman (1991) talks about the agglomeration effect leading to the mushrooming of urban centres. Moomaw and Shatter (1996) explain how Gross Domestic Product (GDP) influences urbanisation in two ways: through the agglomeration benefit that arises due to increasing market size and through the impact of industrialisation. Demurger (2001) shows that infrastructure investment leads to the growth of both urban and rural areas but has a higher impact in urbanised provinces. Xie et al. (2009) examines the short-term and long-term relationship between electricity consumption and urbanisation using an error correction model, the Granger causality test, impulse response, and variance decomposition. They find a steady long-term relationship between the two for China with a feedback effect; i.e., bidirectional causality only in the long run. Hofmann et al. (2013) show the strong influence of per capita GDP

growth, education, and industrialisation on urbanisation, but find no significant effect of road density. Tan et al. (2014) show that all levels of road network have had a considerable effect on the shape and density of the urban landscape of Wuhan, China. Chen (2016) finds road density and distances to transportation services, banks, and hotels to be the most important factors in the urbanisation of the city of Guangzhou. Wan & Zhang (2017) study the role of ICT in facilitating the process of urbanisation and find that information enrichment is important in explaining the worldwide acceleration of urbanisation, with conventional factors losing importance over time. Li (2017) investigates two opposing hypotheses explaining the infrastructure–urbanisation relationship in the Chinese economy for the period 2000–2012: that infrastructure investment drives/does not drive urbanisation. The study uses fixed-effect panel data regressions and the results support the second hypothesis that infrastructure investment pushes urbanisation. The implied turning point of the inverse-U shape relationship was 0.47. In other words, when the urbanisation rate of a Chinese city is lower than 0.47, rising urbanisation is accompanied by rising infrastructure investment intensity in GDP. After the urbanisation rate surpasses 0.47, the infrastructure investment intensity starts to decline due to decreasing demand.

Liddle et al. (2013) show that energy consumption is likely to foster the urbanisation process in two ways: energy and electricity availability improves the quality of life (air conditioning, refrigeration, machinery), and energy consumption is essential for manufacturing to prosper and provide jobs. Wang et al. (2019) find a strong feedback effect between road infrastructure and urbanisation for the Pakistani economy. Shen (2020) differentiates between determinants of state-sponsored and instantaneous urbanisation for the province of Fujian (China) and shows that state-sponsored urbanisation is strongly path-dependent, based on the initial level of urbanisation and development, while spontaneous urbanisation is more dynamic and depends on manufacturing sector expansion. Grekou et al. (2020) explain how FDI inflows significantly influence the momentum of urbanisation on the African continent.

Regarding India, Pandey (1977) attempts to assess the determinants of urbanisation in India and finds that industrialisation positively influences the level of urbanisation, while cropping intensity shows a negative impact.

Surprisingly, the average worker's income has no effect. Sridhar (2005) suggests that the growth centres in India have mushroomed due to infrastructure availability in the form of power, telecommunications, roads, and banking. Pradhan (2007) verifies the existence of a strong relationship between infrastructure and urbanisation in India, with the coefficient of determination concluding that about 27% of the systematic variation in the level of urbanisation is explained by infrastructure availability. Narayana (2011) shows that ICT positively influenced growth in Bangalore by urbanising the city. Tripathi (2017) studies the determinants of urbanisation at the city level and concludes that infrastructure facilities, measured by the number of electricity connections, educational institutions (schools, colleges, and universities), and sanitation facilities promote the pace of urbanisation. For India, Maparu and Mazumder (2017) show that transport infrastructure shared a causality with urbanisation over the period 1991–2017. Hasan et al. (2018) show that cities with a larger share of employment in manufacturing than services tend to grow faster. Diversified manufacturing adds another dimension to the growing urbanisation of cities. Human capital and infrastructure provision (transport and power) within cities fail to produce any significant effect. On the contrary, better connectivity measured by market access systematically affects the growth of cities. Guha (2020) finds that improvements in health infrastructure and energy consumption have led to de-urbanisation across the districts of Assam, while improvements in educational infrastructure and warehousing facilities complemented the growth of urbanisation in the state.

A paper by Liu et al. (2015) deserves special mention. They discuss how in China economic growth affects the pace of urbanisation heterogeneously, with the relationship between economic growth and urbanisation assuming varying patterns. In the northern coastal region and most of inland China, urbanisation Granger causes economic growth, whereas economic growth does not have a significant effect on urbanisation except in the southern coastal and inland regions. The provinces in the southern coastal region do not show a Granger causality relationship between urbanisation and economic growth, thus implying that the effect of economic growth on urbanisation is restricted by administrative intervention.

The two hypotheses we test in this paper are:

H_{01} Infrastructure influences urbanisation uniformly.

H_{02} Infrastructure investment is subject to increasing returns.

3. DATABASE AND METHODOLOGY

This study uses a panel dataset for the 27-year period 1991–2017 for 17[1] Indian states. The states are grouped based on income criteria (Bajar 2013)[2] and are categorised as high-income, middle-income, and low-income. The relationship between the variables is analysed using fixed effects regression. The model is specified in the following form:

$$U_{it} = \alpha + \mu_i + \lambda_t + \beta_1 X + \beta_2 I + \varepsilon_{it} \quad (1)$$

where U_{it} represents the urbanisation rate of state 'i' in time period 't', μ_i represents state fixed effect, λ_t denotes year fixed effect, X includes the control variables based on previous work by Hofmann (2013) and others, and I represents the variable of interest, which in the present case is infrastructure. The list of control variables and the interest variables are specified in Table A2 in the Appendix. The Physical Infrastructure Index (in per cent) comprises teledensity, electricity consumption, and road density; the Social Infrastructure Index (in per cent) comprises Infant Mortality rates and Gross Enrolment Ratio, and ε_{it} is the error term.

The index is constructed based on Principal Component Analysis (PCA), which assigns weights to the broad indicators in an unbiased manner. PCA that is used to compute factor loadings and weights requires that data be unit-free or normalised. After having normalised the data, the index is constructed. The

[1] Andhra Pradesh, Assam, Bihar, Gujarat, Assam, Haryana, Himachal Pradesh, Jammu & Kashmir, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, West Bengal

[2] States are classified as rich if their average PCNSDP is more than India's mean (PCNDP+0.5 (standard deviation)), poor if it is less than India's mean (PCNDP–0.5 (standard deviation)), and middle income if it lies in between.

dimension index formula given by UNDP is used to normalise the data. The normalised values of each variable lie between 0 and 1.

$$\text{Index}=\frac{Actual\,value-Minimum\,value}{Maximum\,value-Minimum\,value} \tag{2}$$

Thereafter the index is prepared using the formula:

$$II=\Sigma W_i X_i/\Sigma W_i \tag{3}$$

where II means infrastructure index (Physical and Social).

After preparing the index it is important to check the time series properties of the individual variables before proceeding, so as to avoid any spurious estimates at the later stage. The results of the unit root tests guide the choice of econometric techniques. For this, the study employs the Augmented Dickey-Fuller (ADF) test using the Levine-Lin-Chu (LLC) (2002) and the Im-Pesaran-Shin (IPS) (2003) methods to check the stationarity properties of the variables. Fisher-ADF, Fisher-PP, and Breitung techniques are also implemented to check the consistency of the results. Here the null of the unit root is tested. The test follows the estimation using the following equation:

$$\Delta Y_t=\alpha_i+\beta_i Y_{it-1}+\sum_{j=1}^{p}\beta_{ij}\Delta Y_{it-j}+\delta_i t+\varepsilon_{it} \tag{4}$$

where i= 1,2,3……N; t= 1, 2 , .T; and Δ is the first difference operator.

4. RESULTS AND DISCUSSION

We first conduct the stationarity tests of our variables so as to avoid spurious results in the analysis. The most popular tests for this available in the literature are in the first-generation testing procedure category. These include IPS (2003) and LLC (2002). Though both of these have different hypotheses they verify the presence or absence of a unit root in the data-generating process. The results from the panel unit root tests are presented in Table 1.

Table 1: Panel Unit Root Statistics

Variable	LLC		IPS	
	Level	Δ	Level	Δ
Ln (Road density)	2.504	−15.336***	5.962	−13.828***
Ln (Rail density)	0.940	−16.455***	2.068	−14.349***
Ln (PCEC)	−0.085	−17.704***	3.383	−17.008***
Ln (Teledensity)	−1.946**	−9.709***	1.799	−9.640***
Ln (PCNSDP)	6.853	−13.941***	12.8081	−14.625***
I/GDP	5.234	0.067	9.928	−3.015***
Infant Mortality Rate	2.998	−5.609***	8.349	−6.522***
Gross Enrolment Ratio in Upper Primary School	1.730	−7.124***	2.713	−9.509***
Agriculture Sector	0.557	−7.912***	4.678	−12.680***
Urbanisation	4.783	−0.664	13.273	−1.368*
PII	1.166	−15.791***	0.919	−17.690***
SII	0.520	−14.433***	−1.572	−15.263***

Note: Model with only constant is adopted.

After checking the stationary properties of the panel data, we report the results arrived at by fixed effects panel regression methodology. In the full sample analysis (Table 2) we find that the indicators infant mortality ratio (IMR) and Enrolment Ratio in Schooling both negatively affect the level of urbanisation in the economy, though the Social Infrastructure Index itself fails to show any significant impact on the dependent variable. The lower IMR value favourably affects our urbanisation level. A lower IMR means that the health of infants is improving. This could be attributed to higher healthcare expenditure by the state government (Barenberg et al. 2017).

On the other hand, the Physical Infrastructure Index shows a positive and significant coefficient. A percentage point improvement in the physical infrastructure is associated with a roughly 0.04% higher urbanisation rate. Unfortunately, social infrastructure does not further the urbanisation level for India as a whole. Positive and significant coefficients for Electricity Consumption

and Teledensity reinstate the promise made by Liddle (2013) and Wan (2017). Liddle identified the importance of electricity consumption for the process of urbanisation while Wan justified the importance of telecommunications for urbanisation by way of reducing information asymmetries. On the other hand, the road infrastructure regressor has a negative and significant coefficient. This could be explained by a better road network (particularly all-weather roads) making travel easier, especially in rural areas where better roads facilitate transportation between home and work and lessen the incentive to relocate. We reject the assertion that growth supports urbanisation, in line with Onjala and Akumu (2016), who in their research reject the credo of growth-led urbanisation in Sub-Saharan Africa, asserting that only developed economies support the hypothesis that growth leads to urbanisation. We find that, overall, in India's urbanisation process physical infrastructure matters more than growth.

The share of the agriculture sector has no significant impact. Another axiomatic finding for the full sample in our case is the validation of an inverse U-effect of infrastructure investment on urbanisation for all 7 regressions. This conclusion is based on the coefficient of the Investment-to-GDP ratio and the square of this ratio represented by Inv Sq. which is in line with Hulten (1994). The coefficient of the linear term is positive (15.75 to 22.05) and ranges between 7.44 and 10.21 for the non-linear term. Both coefficients are significant at the 1% level. Thus, there is no threshold level of Investment to GDP. This means that if the Investment-to-GDP ratio is lower than the threshold level, more investment per GDP will engender urbanisation. However, after the investment/GDP ratio surpasses the threshold level the urbanisation intensity tends to decline.

Table 2: Fixed effects regression model with urbanisation as dependent variable (full sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln PCNSDP	−0.65*	−0.62*	−0.81**	−0.60*	−0.67*	−0.04	−0.79**
	(0.296)	(0.299)	(0.270)	(0.281)	(0.293)	(0.042)	(0.299)
Agri	−0.01	−0.03	0.02	−0.01	−0.04	20.39***	−0.01
	(0.044)	(0.044)	(0.041)	(0.042)	(0.044)	(2.880)	(0.045)
I/GDP	21.73***	22.05***	15.75***	17.02***	19.87***	−0.76**	21.16***
	(2.975)	(2.985)	(2.804)	(2.913)	(2.994)	(0.284)	(3.034)
Inv Sq	−9.80***	−10.21***	−7.56***	−7.75***	−8.91***	−9.28***	−9.65***
	(1.819)	(1.828)	(1.689)	(1.762)	(1.826)	(1.760)	(1.846)
IMR	−0.05**						
	(0.017)						
GER		−0.02**					
		(0.009)					
PCEC			0.01***				
			(0.001)				
TELE				0.07***			
				(0.010)			
Rden					−0.08***		
					(0.002)		
PII						0.04***	
						(0.007)	
SII							0.01
							(0.012)
Constant	30.09***	28.79***	24.80***	26.46***	28.33***	26.54***	26.89***
	(2.163)	(2.031)	(1.827)	(1.880)	(1.967)	(1.903)	(2.067)
N	459	459	459	459	459	459	459
Adj R^2	0.514	0.512	0.590	0.557	0.520	0.546	0.505
AIC	1894.4	1896.1	1816.7	1851.9	1889.1	1862.9	1902.8
F	17.14	17.03	22.75	20.09	17.49	19.30	16.59
Rmse	1.878	1.881	1.726	1.793	1.867	1.815	1.895
T. effect / S. effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Rmse is the Root Mean Square Error, Ln is logarithm, PII is Physical Infrastructure Index, SII is Social Infrastructure Index, Agri is share of agriculture sector in GDP, I/GDP is Investment-to-GDP ratio, Inv Sq is square of I/GDP ratio, IMR is Infant Mortality Rate, GER is Gross Enrolment Ratio in upper primary, Rden is Road density, PCEC is Per Capita Electricity Consumption, TELE is Teledensity.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We move from the full-sample regression to an income-specific regression where the Indian states are grouped as high-income, middle-income, and low-income. Tables 3–5 provide circumstantial evidence on the determinants of urbanisation in each cluster.

Table 3: Fixed effects regression model with urbanisation as dependent variable (High-income states)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln PCNSDP	−1.05	−2.64**	−2.20**	−2.79***	−2.74**	−2.49**	−3.01***
	(0.670)	(0.903)	(0.834)	(0.750)	(0.853)	(0.788)	(0.864)
Agri	−0.13*	−0.05	−0.12	−0.00	−0.16	−0.24**	-0.02
	(0.066)	(0.090)	(0.084)	(0.075)	(0.088)	(0.082)	(0.085)
I/GDP	14.76***	19.15***	19.23***	27.07***	21.92***	29.08***	14.75**
	(3.756)	(5.161)	(4.714)	(4.328)	(4.851)	(4.654)	(5.034)
Inv Sq	−5.06*	−7.44*	−6.66*	−9.86***	−7.01*	−10.16***	−5.46
	(2.124)	(2.907)	(2.677)	(2.414)	(2.739)	(2.547)	(2.812)
IMR	−0.32***						
	(0.027)						
GER		0.01					
		(0.019)					
Rden			−0.00***				
			(0.000)				
PCEC				0.01***			
				(0.001)			
TELE					0.10***		
					(0.027)		
PII						0.08***	
						(0.012)	
SII							0.09***
							(0.024)
Constant	55.78***	43.13***	46.55***	37.87***	47.00***	42.16***	40.30***
	(4.325)	(5.739)	(5.344)	(4.838)	(5.505)	(5.038)	(5.536)
N	216	216	216	216	216	216	216
Adj R^2	0.804	0.644	0.687	0.746	0.671	0.719	0.669
AIC	849.7	978.3	950.5	905.5	960.8	927.0	962.2
F	29.60	13.76	16.43	21.56	15.40	18.97	15.26
Rmse	1.648	2.219	2.081	1.875	2.131	1.971	2.138
T. effect / S. effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Rmse is the Root Mean Square Error, Ln is logarithm, PII is Physical Infrastructure Index, SII is Social Infrastructure Index, Agri is share of agriculture sector

in GDP, I/GDP is Investment-to-GDP ratio, Inv Sq is square of I/GDP ratio, IMR is Infant Mortality Rate, GER is Gross Enrolment Ratio in upper primary, Rden is road density, PCEC is Per Capita Electricity Consumption, TELE is Teledensity.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Fixed effects regression model with urbanisation as dependent variable (Middle-income states)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln PCNSDP	−0.35	−0.29	−0.27	0.49	0.11	0.22	−0.26
	(0.477)	(0.471)	(0.449)	(0.431)	(0.429)	(0.464)	(0.475)
AGRI	0.08	0.11***	0.07*	0.12***	0.14***	0.13***	0.10**
	(0.042)	(0.033)	(0.035)	(0.029)	(0.031)	(0.032)	(0.034)
I/GDP	−1.85	−1.34	0.88	−1.54	−1.15	−2.27	−1.48
	(2.990)	(3.069)	(3.133)	(2.677)	(2.768)	(2.887)	(3.103)
In Sq	0.08	−0.12	−1.63	0.81	0.52	0.68	−0.06
	(1.786)	(1.821)	(1.888)	(1.600)	(1.648)	(1.728)	(1.842)
IMR	0.01						
	(0.013)						
GER		−0.008					
		(0.008)					
Rden			−0.001*				
			(0.0005)				
PCEC				−0.004***			
				(0.0009)			
TELE					−0.07***		
					(0.018)		
PII						−0.02**	
						(0.007)	
SII							−0.006
							(0.008)
Constant	19.63***	20.31***	21.28***	16.29***	16.99***	17.04***	20.14***
	(2.301)	(2.480)	(2.362)	(2.131)	(2.180)	(2.338)	(2.563)
N	135	135	135	135	135	135	135
Adjusted R^2	0.680	0.678	0.694	0.743	0.726	0.701	0.676
AIC	301.4	302.2	295.0	271.7	280.1	292.2	303.0
F	10.29	10.21	10.95	13.61	12.59	11.24	10.14
Rmse	0.681	0.683	0.665	0.610	0.629	0.658	0.685
T. effect/ S. effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Rmse is the Root Mean Square Error, Ln is logarithm, PII is Physical Infrastructure Index, SII is Social Infrastructure Index, Agri is share of agriculture sector in GDP, I/GDP is Investment-to-GDP ratio, Inv Sq is square of I/GDP ratio, IMR is Infant Mortality Rate; GER is Gross Enrolment Ratio in upper primary, Rden is road density, PCEC is Per Capita Electricity Consumption, TELE is Teledensity.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Fixed effects regression model with urbanisation as dependent variable (Low-income states)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln PCNSDP	–0.03	0.17	0.19	0.30	0.35	0.23	0.32
	(0.324)	(0.348)	(0.386)	(0.320)	(0.361)	(0.369)	(0.309)
Agri	0.16**	–0.02	0.04	–0.04	0.05	0.04	–0.01
	(0.055)	(0.060)	(0.056)	(0.053)	(0.054)	(0.056)	(0.048)
I/GDP	46.25***	36.47***	32.92**	19.96*	29.13**	32.60**	46.66***
	(9.69)	(10.16)	(10.78)	(9.61)	(10.37)	(10.51)	(9.28)
Inv Sq	–49.59***	–42.65***	–37.07**	–36.55***	–37.60**	–37.52**	–55.18***
	(10.31)	(11.04)	(11.56)	(9.92)	(11.01)	(11.77)	(10.17)
IMR[3]	0.14***						
	(0.030)						
GER		–0.02*					
		(0.011)					
Rden			0.000				
			(0.000)				
PCEC				0.01***			
				(0.002)			
TELE					0.11*		
					(0.057)		
PII						0.006	
						(0.030)	
SII							–0.06***
							(0.012)
Constant	–0.616	18.78***	14.52***	15.75***	13.66***	14.34***	16.92***
	(4.013)	(3.063)	(2.776)	(2.374)	(2.641)	(2.708)	(2.321)

[3] When the analysis is carried out in all other states using urban mortality rates the results are the same, except for in the low-income states where the coefficient is negative instead of positive.

N	108	108	108	108	108	108	108
Adj R^2	0.378	0.264	0.196	0.380	0.238	0.196	0.423
AIC	319.2	337.4	347.0	318.9	341.2	347.0	311.2
F	3.198	2.336	1.937	3.212	2.175	1.939	3.625
Rmse	0.959	1.044	1.091	0.958	1.062	1.091	0.924
T. effect/S. effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. Rmse is Root Mean Square Error, Ln is logarithm, PII is Physical Infrastructure Index, SII is Social Infrastructure Index, Agri is share of agriculture sector in GDP, I/GDP is Investment-to-GDP ratio, Inv Sq is square of I/GDP ratio, IMR is Infant Mortality Rate; GER is Gross Enrolment Ratio in upper primary, Rden is road density, PCEC is Per Capita Electricity Consumption, TELE is Teledensity.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3 shows the importance of both physical and social infrastructure in high-income Indian states. The regression coefficient tells us that a percentage point improvement in physical and social infrastructure can further urbanisation by 0.08 and 0.09 percentage points respectively. While almost all other infrastructure regressors are equally significant, only the enrolment ratio is largely positive, although not significant. The inverse U-effect of investment on urbanisation remains strong even in the case of high-income states. Agriculture likewise is largely negative, and not at all significant. Thus, the results for the high-income states are as observed for the overall panel.

Switching to the regression results for middle-income states (Table 4) unravels various aspects of the urbanisation determinants. The most significant are:

1. The impact of GDP on the level of urbanisation is insignificant. Also, the results seem to offer plural conclusions. In some regression results the coefficient is positive while in others it is negative.
2. The coefficient on the agriculture regressor is positive and significant. The most likely explanation is the age-old linkage between agriculture and industry that was first talked about by Nurkse and Lewis in balanced growth theory, which posits that linkages between the two economic sectors are more of a norm than an exception.[4] This effect seems to be more prominent in

[4] Refer to Isaksson (2009) for details of the linkage effect between agriculture and industry.

middle-income states because the agriculture sector is under-developed, while most, but not all high-income states predominate agriculturally.

3. The inverted-U hypothesis of infrastructure-led urbanisation is vacuous. This conclusion is mainly drawn on the basis of insignificant values for I-GDP ratio and its quadratic term.
4. Regarding the infrastructure regressors, the results are checkered. Social infrastructure has very little effect on urbanisation. Even the indicators of the social infrastructure index in isolation are extraneous to our dependent variable, while on the other hand the physical infrastructure index is negative and significant at the 1% level. This means that a percentage point improvement in physical infrastructure leads to a decline in urbanisation by 0.02 percentage points. All indicators of physical infrastructure are negative and of a significant size, implying that improvement in any physical infrastructure indicator, be it roads, electricity, or tele density, will decrease urbanisation. Hence, we can say that for middle-income states infrastructure is no less than a 'cog in the wheel'.

Lastly, looking at the regression results for low-income states (Table 5), we find that a percentage point improvement in social infrastructure reduces urbanisation intensity by 0.06 points. The value is highly significant at 0.1%. On the other hand, physical infrastructure has no effect at all, while better teledensity and per capita electricity consumption significantly promote urbanisation. The inverted U-shape of infrastructure investment is explained by the ginormous magnitude of the I/GDP ratio and its square term. Interestingly, the magnitude of the two terms is largest in the low-income states. This implies that the return to investment and urbanisation is highest in the low-income states, presenting an investment opportunity that would stimulate both urbanisation and economic [5]. Similarly, GDP and the share of the agriculture sector in overall GDP are very small.

To sum up, physical infrastructure is an important determinant of urbanisation in high-income states and overall; social infrastructure is important for high-income states and low-income states (where the magnitude is negative); and electricity consumption and teledensity significantly and positively affect urbanisation in high- and low-income states. Road infrastructure in high- and

[5] The infrastructure investment and growth literature is explained in Li (2017).

middle-income states, IMR in high-income states and overall, and enrolment ratio in low-income states and overall affect urbanisation negatively. The supply-led inverted-U hypothesis of infrastructure-led urbanisation is only invalid for middle-income states and holds strongly in all other cases, and most strongly in low-income states. Neither gross domestic product nor share of agriculture in overall GDP significantly affect urbanisation.

In conclusion, the impact of infrastructure on urbanisation across states differs not only according to type of infrastructure but also according to the state's income category. Thus, our hypothesis that infrastructure uniformly influences the level of urbanisation in Indian states is rejected. Regarding the other research hypothesis that infrastructure investment is subject to increasing returns, we strongly reject the null hypothesis, with evidence in favour of decreasing returns. Diminishing returns to infrastructure have also been validated in previous research (Sutherland et al. 2009**;** Canning and Fay 1993; Hulten and Schwab 1993).

The present research follows the direction of Liu et al. (2014), who find that economic growth plays a heterogeneous role in facilitating urbanisation across Chinese regions. Similarly, our analysis investigates the asymmetrical role played by infrastructure investment and infrastructure in influencing urbanisation in Indian states.

5. CONCLUSIONS AND POLICY IMPLICATIONS

The present globalised world is witnessing rapid urbanisation. Some regions urbanise rapidly while others proceed at a slower pace. Because the process is omnipresent, various possible determinants have frequently appeared in the literature. It is imperative to know which are the important factors affecting the urbanisation process in India.

The present paper addresses this concern in a holistic way. Using data on 17 Indian states for 1991–2017, the study aims to discover possible determinants of urbanisation, with special emphasis on infrastructure and infrastructure investment.

The findings suggest various ways forward for policy planners. Relying on the findings of the whole sample can be very misleading because the determinants of urbanisation are not uniform across all Indian states. Because the findings vary between states a 'one policy fits all' approach is inadvisable. The priorities of the states vis-a-vis the urbanisation process will differ.

On the basis of the above research on the determinants of urbanisation in the Indian states, the following recommendations can be made.

1. For high-income states, physical infrastructure (overall), teledensity, electricity consumption, and social infrastructure (overall) positively influence urbanisation, while road network and infant mortality rate have a negative effect. This means that for these states having better infrastructure points to possible future urbanisation and they need to plan accordingly in order to prevent congestion in the later stages of growth, because urban over-crowding can adversely affect human productivity. Any decision regarding infrastructure investment should by default address urbanisation concerns because they work in tandem.
2. For middle-income states, the policy implication is slightly different from that for high income-states. Because in their case the infrastructure components have a negative sign, this does not obviate the need to invest in infrastructure. Rather, it indicates a serious problem with infrastructure quality in these states, which is why even the I/GDP ratio is not significant.
3. For low-income states what is important as of now is improving the level of social services, because the impact of social infrastructure on urbanisation is negative. The reason for this negative impact is the high infant mortality rates of low-income states. On the other hand, because physical infrastructure in isolation positively affects the dependent variable (percentage of urbanisation), the emphasis needs to be on physical infrastructure rather than social infrastructure if these states want more urbanised centres.
4. Lastly, the government needs to be mindful that the infrastructure investment they make in their respective economies has an upper limit beyond which the positive benefits and propensity to urbanise will decline. Thus, ill-considered infrastructure investment will prove wasteful for the economies.

REFERENCES

Angel, S., Jason, P., Daniel, L. C. & Potere, D. (2010). A planet of cities: urban land cover estimates and projections for all countries, 2000–2050. *Working Paper, Lincoln Institute of Land Policy, Cambridge, MA*

Bajar, S (2013). The Infrastructure-Output Nexus: Regional Experience from India. Paper presented at CESifo Conference on The Economics of Infrastructure Provisioning: The Changing Role of the State, at Venice International University, Italy.

Barenberg, A.J., Basu, D. & Soylu, C. (2017). The effect of public health expenditure on infant mortality: Evidence from a panel of Indian states, 1983–1984 to 2011–2012. *The Journal of Development Studies*, 53(10), 1765-1784, DOI: 10.1080/00220388.2016.1241384

Bertinelli, L. & Strobl, E. (2007). Urbanisation, urban concentration and economic development. *Urban studies*, 44(13), 2499–2510 https://doi.org/10.1080/00420980701558442

Black, D. & Henderson, J.V. (1999) A theory of urban growth. *Journal of Political Economy,* 107, 252–284. https://doi.org/10.1086/250060

Bruckner, M. (2011). Economic Growth, Size of the Agricultural Sector, and Urbanization. *The University of Adelaide School of Economics Research Paper No. 2011-16.*

Canning, D., & M. Fay. (1993). The Effect of Transportation Networks on Economic Growth. *Columbia University, mimeo.*

Chakravorty, S. (2007). *Made in India: The Economic Geography and Political Economy of Industrialization*. New Delhi:Oxford University Press.

Chang, G.H. & Brada, J.C. (2006). The paradox of China's growing under-urbanisation. *Economic Systems*, 30, 24–40. https://doi.org/10.1016/j.ecosys.2005.07.002

Chen, Y., Chang, K.T., Han, F., Karacsonyi, D. & Qian, Q. (2016). Investigating urbanization and its spatial determinants in the central districts of Guangzhou, China. *Habitat International*, 51. 59–69.https://doi.org/10.1016/j.habitatint.2015.10.013

Chi, G., Deller, S.C. & Voss, P.V. (2006). Rethinking Highway effects on Population Change. *Public Works Management and Policy,* 11 (1), 18–32. https://doi.org/10.1177/1087724X06292336

Chotia, V & Rao, N. (2015). Examining the Inter-linkages between Regional Infrastructure Disparities, Economic Growth and Poverty: A Case of Indian States. *Economic Annals*, vol. LX (205), 53–66. https://doi.org/10.2298/EKA1505053C

Davis, K & Golden, H. H. (1954). Urbanization and the development of pre-industrial areas. *Economic Development and Cultural Change*, 3 (1), 6–26.

Démurger, S. (2001). Infrastructure development and economic growth: An explanation for regional disparities in China? *Journal of Comparative Economics*, 29 (1), 95–117. https://doi.org/10.1006/jcec.2000.1693

Duranton, G. (2014). Growing through cities in developing countries. *Policy Research Working Paper Series 6818, World Bank*.

O. (2020). Understanding how foreign direct investment inflows impact urbanization in Africa. *International Economics,* 164, 48–68. https://doi.org/10.1016/j.inteco.2020.07.002

Guha, P. (2020). Supply Side Factors in Determining Urbanisation in Assam. A District Level Analysis. In *Understanding Urbanisation in North East India Issues And Challenges,* by Amarjeet Singh and Komol Singha.

Hasan, R., Yiang, J. & Kundu, D. (2018). What made Indian Cities and Towns Grow in the 2000s? Stylized Facts and Determinants. *National Council of Applied Economic Research. India Policy Forum*.

Harris, J.R. & M.P. Todaro (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60, 126–142. https://www.jstor.org/stable/1807860

Henderson, J. V. (2003). The urbanisation process and economic growth: The so-what question. *Journal of Economic Growth*, 8(1), 47–71. https://doi.org/10.1023/A:1022860800744

Hofmann, A., & Wan, G. (2013). Determinants of Urbanisation. *ADB Economics Working Paper Series 355*.

Hulten, C. & Schwab, R. (1993). Infrastructure spending: where do we go from here? *National Tax Journal*, 46, 261–273.

Hulten, C. R. (1994). Optimal Growth with Infrastructure Capital: Theory and Implications for Empirical Modeling. *Working Paper. College Park, MD: University of Maryland*.

Im, K.S., M. H. Pesaran, & Y. Shin. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*. 115, 53–74. http://dx.doi.org/10.1016/S0304-4076(03)00092-7

Isaksson, A. (2009). Public Capital, Infrastructure and Industrial Development. *United Nations Industrial Development Organization Working Paper 12/2009.*

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99 (3), 483–499. https://doi.org/10.1086/261763

Leitão, N. (2013). A panel data modelling of agglomeration and growth: cross-country evidence. *Theoretical and Empirical Researchers in Urban Management*, 8 (1), 67–77.

Levine, A., C. F. Lin,& C. S. Chu. (2002). Unit root tests in panel data: Asymptotic and finite sample properties. *Journal of Econometrics,* 108, . 1–24. https://doi.org/10.1016/S0304-4076(01)00098-7

Liu, T.Y., Su, C.W. & Jiang, X.Z. (2015). Is economic growth improving urbanisation? A cross-regional study of China. *Urban Studies*, 52(10), 1883–1898. https://doi.org/10.1177/0042098014540348

Li, Z. (2017). Infrastructure and Urbanization in the People's Republic of China. *Asian Development Bank Institute* (Issue 632).

Liddle, B & Lung, S. (2013). Might electricity consumption cause urbanization instead? Evidence from heterogeneous panel long-run causality tests. *MPRA Paper No. 52333.* https://mpra.ub.uni-muenchen.de/52333/

Maparu, T.S & Majumder, T.N. (2017). Transport infrastructure, economic development and urbanization in India (1990–2011): Is there any causal relationship? *Transportation Research* Part A, 100, 319–336. https://doi.org/10.1016/j.tra.2017.04.033

Moomaw, R. & Shatter, A. M. (1996). Urbanization and economic development: A bias toward large cities? *Journal of Urban Economics*, 40(1), 13–37. https://doi.org/10.1006/juec.1996.0021

Narayana, M.R. (2011). Globalisation and urban economic growth: Evidence from Bangalore (India). *International Journal of Urban and Regional Research*, 35, 284–301.

Onjala, J. & K'Akumu, O.A. (2016). Relational patterns of urbanisation and economic growth in sub-Saharan Africa. *Development Southern Africa*, 33(2), 234–246. DOI: 10.1080/0376835X.2015.1120655.

Pandey, S.M. (1977). Nature and determinants of urbanization in a developing economy: The case of India. *Economic Development and Cultural Change*, l 25 (2), 265–278. DOI: 10.1086/450946

Pradhan, R. P. (2007). Does infrastructure play a role in urbanization? Evidence from India. *Indian Journal of Economics & Business*, 6(1).

Quintana**,** D.C. & Royuela V. (2014). Agglomeration, inequality and economic growth. *Annals of Regional Science*, 52 (2), 343–366. https://doi.org/10.1007/s00168-014-0589-1

Quintana, D.C. (2017). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics,*

Vol 98, 158–173. https://doi.org/10.1016/j.jue.2016.02.003

Sekkat, K. (2013). Does infrastructure mitigate the effect of urban concentration on poverty in developing countries? Economic Research Forum Working Paper Series, WP No. 800.

Sridhar, K.S. (2005). *Incentives for Regional Development: Competition among Sub-National Governments*. Palgrave Macmillan: Basingstoke.

Sutherland, D., Araujo, S., Égert, B., & Kozluk, T. J. (2009). Infrastructure investment: links to growth and the role of public policies.

Tan, R., Yaolin, L., Yangfang, L., Qingsong, H., Licai, M. & Shuohua, T. (2014). Urban growth and its determinants across the Wuhan urban agglomeration, central China. *Habitat International,* 44, 268–281. https://doi.org/10.1016/j.habitatint.2014.07.005

Tiebout, C. M. (1956). A pure theory of local expenditures. *The Journal of Political Economy*, 64(5), 416-424

Todaro, M.P. (1969). A model of labor migration and urban unemployment in less developed countries. *American Economic Review*, 59, 138–148. https://www.jstor.org/stable/1811100

Tripathi, S. (2017). Relationship between infrastructure and population agglomeration in urban India: An empirical assessment. ADBI Working Paper 731, Tokyo: Asian Development Bank Institute.

Tripathi, S. & Rani, C. (2018). Determinants of urbanisation in different classes of cities and towns in India. *Review of Development & Change*, XXIII (1), 121–148.

Turok, I. & Gordon, M.G. (2013). Urbanization and economic growth: the arguments and evidence for Africa and Asia. *Environment & Urbanization*, 25(2), 465–482. DOI: 10.1177/0956247813490908.

Wan, G., & Zhang, Y. (2017). *Accelerating Urbanization Explained: The Role of Information. Asian Development Bank Institute, 674.*

Zhang, K. H. (2002). What explains China's rising urbanisation in the reform era? *Urban Studies*, (12), 2301–2315. https://doi.org/10.1080/0042098022000033881

Received: February 04, 2021
Accepted: February 28, 2022

APPENDIX

A1: List of states according to income bracket

High-income states	Middle-income states	Low-income states
Gujarat	Andhra Pradesh	Assam
Haryana	Jammu and Kashmir	Bihar
Himachal Pradesh	Odisha	Madhya Pradesh
Karnataka	Rajasthan	Uttar Pradesh
Kerala	West Bengal	
Maharashtra		
Punjab		
Tamil Nadu		

Source: Author's Compilation

A2: Data sources of the variables used in the study

Variable	Data Source
• Investment (Capital Expenditure + Outstanding Credit in Scheduled Commercial Banks)	RBI Handbook of Statistics on Indian States
• Agriculture Sector (Share of agriculture in overall NSDP) (in per cent)	Computed using data on overall GDP and Sectoral GDP
• Per Capita Net State Domestic Product (Expressed in 2011–12 prices) (log form)	RBI Handbook of Statistics on Indian States
• Infant Mortality Rate	Sample Registration System Bulletins
• Gross Enrolment Ratio in Upper Primary School	Economic Survey Series of India
• Road Density (per 1000 sq. km of geographical area)	Handbook of Statistics on Indian States
• Per Capita Electricity Consumption (In KWh)	Handbook of Statistics on Indian States
• Teledensity (Per 100 population)	Handbook of Statistics on Indian States

• Urbanisation	The value is interpolated for 1991–2000. Post 2000 the population projections are taken from the Report of the Technical Group on Population Projections Constituted by the National Commission on Population.

Source: Author's Compilation

ECONOMIC ANNALS, Volume LXVII, No. 232 / January – March 2022
UDC: 3.33 ISSN: 0013-3264

https://doi.org/10.2298/EKA2232105T

*Dragan Tevdovski**
*Petar Jolakoski***
*Viktor Stojkoski****

DETERMINANTS OF BUDGET DEFICITS: THE EFFECTS OF THE COVID-19 CRISIS

ABSTRACT: *This paper revisits the discussion on the determinants of budget balances and investigates the change in their effect in the context of the COVID-19 crisis. The analysis uses data on 43 countries and a system generalised method of moments approach. The results show that the overall impact of the global pandemic has led to a disproportionate increase in the estimated effects of the macroeconomic determinants on the budget balance. We also find that more developed economies were able to implement higher stimulus packages for the same relative level of primary balance. We believe that one of the factors affecting this outcome is that more of their government debt is held in domestic currency.*

KEY WORDS: *budget deficits, economic determinants, COVID-19*

JEL CLASSIFICATION: H11, H62, I18

* Faculty of Economics, University Ss. Cyril and Methodius in Skopje, Skopje, North Macedonia
** Association for Research and Analysis – ZMAI, Skopje, North Macedonia
*** Corresponding author; Faculty of Economics, University Ss. Cyril and Methodius in Skopje, Skopje, North Macedonia; Association for Research and Analysis – ZMAI, Skopje, North Macedonia; vstojkoski@eccf.ukim.edu.mk

1. INTRODUCTION

Governments around the world increased budget deficits in response to the sharp decline in global output that resulted from the extensive economic disruption caused by the COVID-19 pandemic. To reduce the impact of the pandemic, governments provided liquidity support and replaced lost household income, saving jobs and preventing large-scale bankruptcies. However, these measures were costly. Together with the drastic fall in tax revenue and the abrupt rise in government expenditure, the measures are expected to push global public debt to an all-time high. The International Monetary Fund (IMF) estimated that global public debt would increase by 16 percentage points in 2020, from 83% of GDP in 2019 to around 100% in 2020, the largest increase ever.

Pre-pandemic literature examined extensively the responsiveness of fiscal policy to the business cycle (for example, Alesina et al., 2008; and Frankel et al., 2013) and showed how political and institutional arrangements affect budget deficits (for example, Roubini and Sachs 1989; Perotti and Kontopoulos, 2002 and Agnello and Sousa, 2009). The focus of this paper is to review and revisit the determinants of budget balances and show how the COVID-19 pandemic has changed their impact. We attempt to answer several compelling questions: whether the effect of increased government expenditure on the budget balance changed in 2020, how the state of the labour market in 2020 affected the budget balance, whether budget balances in 2020 were constrained by the prevailing debt levels and current low, long-term interest rates, and whether the existence of previous vulnerabilities mattered for the change in the 2020 budget balance.

To provide the answers, we use a system-GMM estimation procedure where we introduce interaction terms for macroeconomic variables and a dummy variable for 2020 in an otherwise standard specification including macroeconomic, political/institutional, and demographic variables, as suggested by the literature (Alesina et al., 1998; Agnello and Sousa, 2009; Maltritz and Wuste, 2015). We use a dynamic panel dataset with data for 43 countries (all countries for which sufficient data is available) for 26 years, from 1995 to 2020. We use actual realisations of all the series from 1995 to 2019 and IMF WEO forecasts of the series for 2020. The system approach allows us to address and overcome the issue of endogeneity that arises among the independent variables and which is not accounted for in a standard least squares regression.

We contribute to the literature in several ways. First, the aim of a growing body of literature is to determine the size of the 2020 budget deficits and suggest policies to repay the increased debt sustainably (Makin and Layton, 2021). However, to the best of our knowledge, the literature has not investigated the effect of economic determinants on the size of the 2020 budget deficits. This is one of the first papers to quantify and analyse the difference in the effect of standard macroeconomic determinants on the budget balance prior to and during the COVID-19 pandemic. In addition, the paper discusses at length the implications of this change in terms of the long-term economic scarring that might result from the on-going crisis.

Our results indicate that the overall impact of the global pandemic has led to a disproportionate increase in the estimated effects of several macroeconomic determinants on the budget balance. In particular, the absolute effects of the rise in government debt and government expenditure on the budget balance were greater in 2020 than in the preceding period. In 2020, higher public debt implied even lower budget deficits, suggesting that higher debt during a severe economic downturn constrained additional government spending.

We also study the relationship between the size of the primary balance in 2020, the fiscal packages introduced to tackle the consequences of the COVID-19 pandemic, and the government debt position by currency denomination. We argue that more-developed economies were able to implement higher stimulus packages for the relatively same level of primary balance, and this was in part because they hold more government debt in domestic currency (as a percentage of GDP).

The remainder of the paper is organised as follows. Section 2 offers a brief overview of the related literature. Section 3 introduces in more detail the applied methodology and the data used. Section 4 provides the empirical results and studies the thereby induced implications. In Section 5 we discuss our findings.

2. LITERATURE REVIEW

The existing literature studies the budget balance from two main perspectives: economic and political. Most studies that focus on the economic perspective analyse the response of fiscal policy to output (Gali and Perotti, 2003; Akitoby et

al., 2004; Talvi and Vegh, 2005; Frankel et al., 2013). Standard Keynesian models posit that fiscal policy should be countercyclical, allowing for budget deficits in recessions and saving for budget surpluses in booms; i.e., government spending (taxes) should rise (decrease) in recessions and vice versa in booms. The tax smoothing theory of Barro (1979) argues that the government should smooth both tax rates and government spending by borrowing during recessions and repaying during booms. However, the literature finds mixed evidence regarding the counter-cyclicality of fiscal policy. The consensus from the studies is that fiscal policy is counter-cyclical in most developed countries, while it is pro-cyclical in developing countries (Afonso et al., 2010).

Explanations of the cross-country variation in fiscal policy cyclicality cannot be solely economic. Governments' political characteristics, ideological motivation, electoral system, and institutional arrangements are also important determinants of fiscal policy. Much of the literature focused on political determinants finds that governments that are weaker in terms of tenure and political power create larger budget deficits. Roubini and Sachs (1989) show that countries where governments have short tenures tend to have higher deficits on average. Moreover, the paper shows that multi-party coalition governments have a greater tendency to develop large and persistent deficits than majority-party governments. Similarly, Lane (2003) argues that countries with dispersed political power are most likely to run pro-cyclical fiscal policies. Higher deficits are also found to be positively associated with the size of the government cabinet, measured as the number of spending ministers (Volkerink and De Haan, 2001; Perotti and Kontopoulos, 2002).

Although it is reasonable to expect that right-wing governments will practice tight fiscal policy and left-wing governments loose fiscal policy, the empirical literature finds mixed evidence for the influence of the government's ideological preferences on the budget balance (Alesina et al., 1997; Mulas-Granados, 2003). Opportunistic governments without ideological preferences that follow policies to maximize their probability of winning the next election tend to have higher budget deficits in election years (Franzese Jr, 2002; De Haan and Mink, 2005). Alesina and Perotti (1995) find that large deficits are more common in countries with proportional rather than majoritarian and presidential electoral systems. In

addition, Alesina et al. (2008) argue that most of the pro-cyclicality of fiscal policy in developing countries can be explained by high corruption levels.

Institutional factors are also positively associated with fiscal performance. Leachman et al. (2007) find that fiscal performance is better when fiscal budgeting institutions are strong. De Haan and Sturm (1997) find that a strong finance minister or a commitment to negotiated budget targets can be especially effective in keeping deficits down in countries where there is political instability. Henisz (2004) finds that checks and balances that limit the discretion of policymakers reduce the volatility of government expenditure and revenue.

Lastly, a relatively new but important strand of literature examines fiscal persistence: the degree to which current fiscal behaviour depends on its own past development. For instance, Afonso et al. (2010) find that countries with higher fiscal persistence tend to have lower discretion.

3. METHODOLOGY

3.1. Model

To construct the econometric model we follow the literature described in the previous section, and specify it as:

$$Balance_{it} = \beta_0 Balance_{it-1} + Y_{it}^{'}\beta_1 + X_{it}^{'}\beta_2 + Y_{it}^{'}D\delta_{\beta_1} + X_{it}^{'}D\delta_{\beta_2} + \varepsilon_{it},$$

where the dependent variable $Balance_{it}$ is the primary budget balance of country i in time t. We assume that the primary budget balance is dependent on its previous value $Balance_{it-1}$, with β_0 being its marginal effect and two disjoint sets of explanatory variables $Y_{it}^{'}$ and $X_{it}^{'}$ and a random error ε_{it}. The first set of explanatory variables has a marginal effect β_1 and describes potential macroeconomic determinants of the primary budget balance, whereas the second set is constituted of political and institutional variables, and β_2 is their marginal effect. In addition, we include a country-specific effect in the equation to account for potentially omitted variables that are invariant over time.

Alesina et al. (1998), Agnello and Sousa (2009), and Maltritz and Wuste (2015) use similar model specifications to understand the critical factors that drive the magnitude and characteristics of budget deficits. The novelty in our specification is the presence of the interaction term between the independent variables and a dummy variable D for 2020, used to quantify the potential different effect of the variables due to the coronavirus pandemic. Concretely, δ_{β_1} and δ_{β_2} represent a direct measure of the change in the effect of the macroeconomic and political and institutional variables respectively. A significant and negative value of δ_{β_1} (δ_{β_2}) implies that the variables' contribution to the size of the budget deficit in 2020 was larger than usual.

A general problem of this model specification is the presence of endogeneity due to potential interdependence between the explanatory variables and the budget balance (Agnello and Sousa 2009), which may lead to biased and inefficient parameter estimates. To account for this problem, we resort to a system GMM parameter estimation. The system GMM solves the endogeneity problem in two steps. In the first step, each variable is first-differenced, thus removing the potential endogeneity due to correlation between the country-specific effects and the explanatory variables (Arellano and Bond 1991). In the second step, the endogeneity between the dependent and explanatory variables is removed by instrumenting the differenced variables with their available lags in levels: the levels of the dependent variable lagged for two or more periods and the levels of the explanatory variables lagged for one or more periods (Blundell and Bond 1998). Another advantage that system GMM has over other estimation procedures is that it efficiently accounts for the potential problem of the lagged dependent variables being weak instruments. Weak instruments cannot be used to solve the endogeneity problem. They usually appear in situations when there is high heterogeneity in the cross-sectional sample or when the time series for each cross-section has a large variance. For example, we are looking at all available country data and do not restrict our sample to specific countries. Hence, there might be high heterogeneity in the cross-sectional sample. Unlike other estimation procedures, system GMM resolves this issue by simply adding additional restrictions to the moment conditions. (See Roodman (2009) for a more detailed explanation of the properties of system GMM.)

3.2. Data

For our analysis we construct an unbalanced panel annual dataset with macroeconomic, political, and institutional variables for 43 countries, with yearly data covering the period 1995 to 2020. Macroeconomic data comes from the IMF World Economic Outlook (WEO) October 2020 database, and political and institutional data from Databanks International's Cross-National Time-Series Data Archive and Polity IV Database. Our starting point was all of the countries included in the WEO database, but some of the countries were removed from the initial sample due to lack of sufficient data availability in other databases.

For the period 1995 to 2019 we use the actual realisations of the chosen variables. However, for 2020 we use forecasts for the macroeconomic variables (IMF) and assume that no change has taken place for the institutional and political variables (i.e., $\delta_{\beta_2} \approx 0$)) . Although this is a rather strong assumption, our rationale is that changes in these variables require a multitude of legislative and political actions that can rarely be achieved in the span of a year. A detailed analysis of all the countries showed that it took at least two years for their polity score to change, and it remained at the same value during the last few years of our sample. Furthermore, for the entire dataset, the type-of-regime variable had little or no variation over the years, and across the majority of countries the size of the cabinet in the last period changed by only one or two ministers.

In all specifications, the dependent variable is the general government primary budget balance as a percentage of GDP. We use this measure because it better matches the discretionary decisions of the fiscal authorities than the overall budget balance. It does not include interest payments for outstanding debt piled up from the previous period, which is not relevant to our study (Maltritz and Wuste, 2015).

The two sets of explanatory variables are based on the relevant literature, described in the previous section. We use a standard set of macroeconomic variables in the first set: lagged primary budget balance, government expenditure (log), gross debt (log), interest rate of government debt securities, unemployment rate, GDP growth rate, and population (log).

For the second set of political and institutional explanatory variables we use polity scale, type of regime, and cabinet size. Polity scale is a variable that evaluates how democratic a country is on a scale from –10 to 10, where the two extremes imply that the country is either fully autocratic (–10) or fully democratic (10). Type of regime is a categorical variable that provides an estimate for the type of government regime in the country: 1) civilian, 2) military-civilian, and 3) military. Size of cabinet quantifies the number of ministers in a government. The list of all used variables, their transformation and data sources are presented in Table A1 in the Appendix.

Table 1 summarizes the descriptive statistics for all of the variables included in our empirical analysis for 2019 and 2020. These statistics suggest that the primary budget deficit across all countries increased by 6.1 percentage points on average between 2019 (–0.32% of GDP) and 2020 (–6.4% of GDP), while gross government debt increased by 12.8 percentage points on average between 2019 (69.3% of GDP) and 2020 (82.1% of GDP). This extraordinary increase in global public debt happened at a time of an almost 1pp decrease in the average interest rates of government securities between 2019 (3.4%) and 2020 (2.5%). During the 2020 crisis, as expected, the loose fiscal policy led to an average increase in government expenditure of 8.7 percentage points, whereas the average increase in the unemployment rate was 2.5 percentage points and the average decrease in the GDP growth rate was 8.6 percentage points across the entire sample. These numbers reflect the current economic conditions and the economic support packages implemented by policymakers around the world as a response to the global pandemic.

The polity scale variable shows that there is more democracy in advanced than in emerging economies (9.1 and 5.3 respectively). The average number of ministers in the governments of both country groups is almost the same (around 21). Most governments in our sample are classified as civilian, with a few exceptions – Algeria, Egypt, Fiji, Sudan, and Thailand are classified as military-civilian, and Pakistan and Thailand have had a military regime for a relatively short period of time.

Table 1: Summary statistics.

	All countries	
Variable	2019	2020
Dependent variable		
Primary balance	–0.32	–6.44
	(2.85)	(4.24)
Macroeconomic and demographic variables		
Gov. gross debt	69.27	82.14
	(38.48)	(46.78)
Population	46.42	47.2
	(73.33)	(77.73)
Unemployment rate	6.78	9.24
	(6.37)	(8.31)
Interest rate	3.41	2.52
	(3.43)	(3.3)
Gov. exp. (% of GDP)	34.77	42.88
	(9.82)	(11.45)
GDP growth rate	2.47	–6.11
	(1.68)	(2.80)
Political and institutional variables		
Polity	7.57	7.4
	(4.08)	(4.37)
Size of cabinet	19.57	19
	(5.51)	(5.47)
Civilian regime	25	21
Military-civilian regime	1	1
Military regime	0	0

Note: Mean values per country group for 2019 and 2020. Standard deviations in brackets. Regime statistics refer to number of countries.

Spending and revenue discretionary budget measures taken to combat the virus by emerging and developing economies account for more than 3.5% of GDP, and more than 9% of GDP in advanced economies (IMF WEO October 2020). Given the specific nature of this shock, the severe weakening of aggregate demand and continuous disruptions of aggregate supply are expected to lead to the deepest global recession since World War II (World Bank Global Economic Prospects

2020). As a result, governments across the globe stepped in with extensive fiscal packages along with complimentary institutions to help the ailing economies, including wage subsidies, tax deferrals, easing of regulatory burdens, transfers to businesses and households, postponement of loan repayments, and government guarantees. Although the economic support offered by all countries is unprecedented and higher than that offered during the global financial crisis in 2009 (World Bank Global Economic Prospects), the state of the economies prior to the pandemic was also a crucial determinant of the magnitude of the response.

4. EMPIRICAL RESULTS

In this section we discuss the empirical results obtained using the Blundell and Bond (1998) methods of implementing a dynamic linear GMM estimation. A summary of our main findings can be found in Table 2. In column 1 we present the results from our model with the primary budget balance and the set of macroeconomic variables. We add additional explanatory variables that have been used in the literature, which include a demographic effect (column 2) and a potential effect resulting from political/institutional variables (column 3). Finally, in column 4 we present results that include interaction terms with the variables that we consider to have been affected by the Covid-19 crisis and a dummy variable for 2020. The last two rows of Table 2 report the results of two statistical tests that evaluate whether our regressions satisfy the baseline regression specification assumption. The first is the AR(2) test, which under the null hypothesis assumes that there is no autocorrelation in the random errors. The second is the Sargan test for overidentification of the instruments used for estimation under the null hypothesis that the over-identifying restrictions are valid. In each case, the statistical tests do not reject the null hypotheses, suggesting that our models are correctly specified.

In all of the regressions we estimated, most of the macroeconomic variables are significant and have the expected sign, as typically found in the literature (Roubini and Sachs, 1989; Bayar and Smeets, 2009; Maltritz and Wste, 2015). In every regression the effect of the lagged primary budget balance is positive and significant. This persistence in the effect of the primary budget balance corresponds to a well-documented inertia in the budgetary process found in the literature. Similarly, the effect of the GDP growth rate is positive and significant – as expected, because when the economy has a higher growth rate the primary

budgetary balance improves in the short run. The estimated coefficient of the unemployment rate is negative and significant in every case, except in the model in column 4 of Table 2 where it loses significance. This implies that an increase in the unemployment rate worsens the primary budgetary balance. It can be argued that this is a result of the additional government expenditure incurred to support the labour market. The stock of debt and the long-term interest rate have a significant and positive relationship with the primary budgetary balance. The fact that gross government debt has a positive effect on the primary budget balance confirms a previous finding in the empirical literature that higher debt improves the primary budget balance and reduces deficits. Maltritz and Wuste (2015) argue that high debt implies less fiscal space to encourage additional spending, while low debt levels enable countries to run higher deficits. We emphasize that higher interest payments from debt do not affect the primary balance, which by definition excludes interest payments. However, higher long-term interest rates on debt instruments also implies less fiscal space and leads to improved budgetary balances.

Next, we observe that government expenditure has a significant and negative impact on the primary budgetary balance, as higher government expenditure raises the primary budget deficit in the short run. Similarly, population size, as a demographic variable, has a negative relationship with primary budget balance in one regression (column 3). In this context, Furceri and Poplawski (2008) argue that a larger pool of taxpayers can be insurance against idiosyncratic shocks, which leads to lower budget deficits. Indeed, this result suggests that the larger countries have greater fiscal space for discretionary actions in a global crisis.

Table 2: Blundell–Bond linear dynamic panel-data estimation

	(1)	(2)	(3)	(4)
VARIABLE	Macroeconomic	Demographic	Political/Institutional	2020
Primary balance (t-1)	0.404***	0.399***	0.374***	0.356***
	(0.018)	(0.026)	(0.020)	(0.043)
GDP growth rate	0.244***	0.251***	0.223***	0.169***
	(0.013)	(0.011)	(0.015)	(0.018)
Unemployment rate	–0.064***	–0.067***	–0.081***	–0.012
	(0.022)	(0.019)	(0.030)	(0.046)
Government debt	2.828***	2.513***	2.830***	3.520***
	(0.376)	(0.319)	(0.337)	(0.527)
Government expenditure	–17.267***	–17.070***	–19.745***	–20.122***

	(0.839)	(1.074)	(1.709)	(2.732)
Interest rate	0.263***	0.243***	0.198***	0.191***
	(0.025)	(0.019)	(0.026)	(0.051)
Population		–0.219	–6.323**	–1.031
		(1.707)	(2.994)	(3.836)
Polity			1.052***	0.673***
			(0.250)	(0.230)
Size of cabinet			–0.032	0.048
			(0.031)	(0.037)
Type of regime (2)			8.901***	8.914**
			(1.318)	(4.142)
Type of regime (3)			–5.435***	1.464
			(3.139)	(2.690)
GDP*2020				0.795
				(0.685)
Unemployment*2020				0.741
				(0.682)
Government expenditure*2020				–6.814***
				(1.898)
Debt*2020				6.781***
				(2.515)
Population*2020				–0.618
				(0.537)
Interest rate*2020				–0.213
				(0.586)
Constant	49.516***	50.961***	67.588***	53.406***
	(2.655)	(5.614)	(10.078)	(11.963)
Observations	854	854	854	854
Number of countries	43	43	43	43
Arellano-Bond test (H0: no autocorrelation of order 2)	p = 0.5358	p = 0.5066	p = 0.2522	p = 0.565
Sargan-Hansen test (H0: overidentifying restrictions are valid)	p = 1.000	p = 1.000	p = 1.000	p = 1.000

Note: Standard errors in parentheses

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

Some of the results for the set of political/institutional variables agree with empirical findings in the literature. The coefficient in front of the polity scale variable is significant and positive, implying that a more democratic regime tends to have stronger institutions and functional checks and balances that limit policymakers' discretion to increase the budget deficit. This finding is in line with

the literature (for example, Henisz 2004 and Leachman et al. 2007). The effect of the cabinet size is negative but insignificant. A negative impact fits the stylized fact in the literature that a higher number of spending ministers is associated with a lower primary balance (for example, Volkerink and De Haan 2001 and Perotti and Kontopoulos 2002). Finally, in terms of the type of regime, the results show that countries that have a military-civilian or civilian regime tend to have higher primary budget deficits on average.

The last column shows the change in the effect of the macroeconomic variables on the primary budget balance in the first year of the global pandemic. It appears that the pandemic increased in absolute values the magnitude of the estimated effects of all of the variables, except the unemployment rate, where the direction of the relationship is reversed. However, the change of the effect in some of the variables is statistically insignificant. This is the case for the GDP, unemployment, population, and interest rate variables. On the other hand, the sharp decline in economic activity led to increased government expenditure, which in turn resulted in higher budget deficits. This is as expected, since all of the countries in our sample implemented some form of economic support package to mitigate the economic cost of the crisis, resulting in a deterioration in their fiscal health. In addition, the resulting increase in public debt leads to a larger positive effect of gross debt on the primary balance in 2020. This can be an indication that higher debt during a severe economic downturn enhances incentives against spending.

The robustness of our results is confirmed by re-estimating the model when the interest rate is removed from the list of explanatory variables. We explicitly choose this variable since, when excluded, the sample size increases by the largest margin. The results are given in Table A2 in the Appendix, which shows that the estimated coefficients are not much different from those presented in Table 2. Therefore, it can be argued that our results are robust.

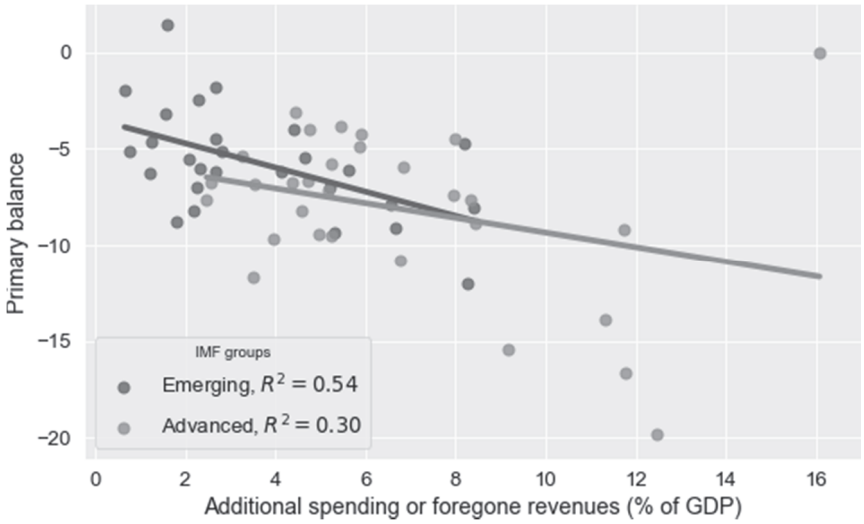
4.1. Implications

In the period ahead, economies globally are expected to experience unprecedented economic and social costs resulting from the COVID-19 pandemic. Returning to pre-crisis levels of economic activity is a daunting task for countries all over the world and the path to recovery will not be smooth, even, or certain.

To begin with, the fiscal stimulus packages implemented by governments in response to the pandemic put a significant strain on public finances. The unparalleled fiscal response to the demand slumps and supply interruptions that followed the crisis was largely implemented in advanced and some large emerging market economies, because they could rely on more favourable financing conditions prior to the crisis and retained the ability to borrow at lower interest rates (IMF, WEO 2020, and IMF Fiscal Monitor 2020).

To better understand the relationship between the fiscal stimulus packages introduced by governments and 2020 budget deficits, in Figure 1 we plot the primary balance in 2020 as a function of the additional spending or foregone revenues in response to COVID-19. The source of data for countries' fiscal measures in response to pandemic is the IMF Fiscal Monitor database. We divide the countries into emerging and advanced based on the IMF's classification.

Figure 1: IMF country groups: Explained variation in budget deficits due to additional spending or foregone revenues in response to COVID-19 pandemic

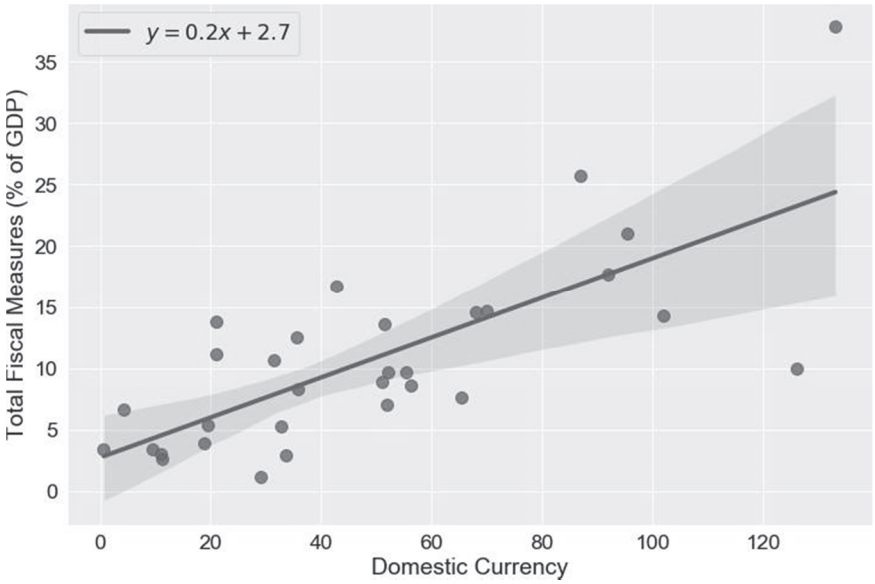


Source: Fiscal Monitor, Database of Country Fiscal Measures in Response to the COVID-19 Pandemic, IMF.

The figure shows that some advanced economies were able to implement higher stimulus packages for the relatively same level of primary balance. The conclusion

is similar when looking at the total fiscal measures undertaken, which include liquidity support in the form of equity injection, loans for asset purchase, and debt assumption. The figure highlights that the total fiscal measures were in general higher in advanced economies than in emerging market economies.

Figure 2: Explained variation in total fiscal measures due to government debt position denominated in domestic currency (as a percentage of GDP)



While country income per capita is an important determinant of the size of total fiscal measures during the pandemic, we want to emphasize that government debt position denominated in domestic currency was also a determinant during the COVID-19 crisis. We could not include this variable in our regression model due to lack of data availability for a lot of the countries in our sample. However, Table A3 in the Appendix presents all of the countries for which data is available in the Fiscal Monitor of the IMF database and the Quarterly Public Sector Debt database of the World Bank on gross central government debt position by currency denomination (as a percentage of GDP) in the first quarter of 2020. We illustrate the relationship using these countries in Figure 2. The figure shows that countries with a higher government debt position denominated in domestic currency (as a percentage of GDP) were able to execute higher total fiscal measures (as a percentage of GDP) during the pandemic. For example, 4 of the 5 countries with

the highest total fiscal measures as a percentage of GDP are Italy, the United Kingdom, France, and Spain, whose debt position in domestic currency is higher than 80%. Also, we note that the correlation coefficient between debt position denominated in domestic currency and size of total fiscal measures is fairly high, 0.737, and significant at the 1% level.

5. CONCLUSION AND DISCUSSION

Fiscal packages to stimulate the economy during the COVID-19 pandemic led to unprecedented growth in budget deficits in almost every country in the world. We analysed whether this growth was also affected by changes in the economic determinants. Using System GMM and data for 43 countries over a period of 26 years, we provided evidence that the growth in budget deficits may have been accelerated by these changes. We found that in the first year of the pandemic the marginal negative impact of government expenditures increased, and the positive impact of budget deficits increased.

We then postulated that these changes resulted from more-developed economies implementing higher stimulus packages for the same level of budget deficit, mainly because of the advantage of servicing their debt in their national currency. While the fiscal packages played a vital role in the governments' efforts to combat the consequences of the pandemic, we also believe that future fiscal space will be limited because of these efforts. One limitation of our analysis is that the study was based on data that captures the impact of the first year of the pandemic. The increase in the global debt-to-GDP ratio will certainly pose additional challenges for debt sustainability of all economies globally in the medium to long run. Equally relevant is the potential risk on the horizon for debt financing conditions. Moreover, some of the vulnerabilities that existed prior to the crisis, such as population ageing, are likely to further contribute negatively to the outlook for the stock of sovereign debt. For instance, in their empirical study of OECD economies, Honda and Miyamoto (2020) find that population ageing weakens fiscal spending effects, and in order to support the economy in a downturn, countries will need to revert to larger fiscal support packages. Finally, the build-up of debt is also expected to constrain future government spending on growth and development, as a large part of government revenues will be consumed by debt service. We argue that in the absence of more refined data, the analysis performed here provides a starting point for the development of a more

comprehensive understanding on how the economic determinants of the budget balance are changing because of the pandemic. We believe that the insights provided by this analysis and an improved understanding of economies' fiscal behaviour during the coronavirus pandemic will aid the development of studies on the long-term impact of COVID-19 on the budget balance, as soon as such data is available.

Last but not least, the size, distribution, and adjustment of the 2020 budget balance will determine both the general social prospects and the economic capabilities of every country in the aftermath of the pandemic (Stojkoski et al., 2020a, b; Tevdovski et al., 2021). It is expected that the structural changes imposed on economies by health authority constraints (social distancing, teleworking, movement restrictions, capacity restrictions) will redistribute societal resources from highly inflexible sectors of the economy to highly adaptive sectors. As a result, welfare losses from labour market distortions (jobs at the lower quantiles of the wage distribution, in informal employment, with temporary working arrangements) and human capital accumulation disruptions (schooling interruptions and transformation) will probably worsen the level of poverty and income inequality worldwide. We believe that in order to tackle the challenges induced by COVID-19, vital mechanisms for restoring the fiscal health of governments will have to be developed. This is the subject of our ongoing research.

REFERENCES

Agnello, L., & Sousa, R. M. (2009). The determinants of public deficit volatility. *ECB working paper.*

Akitoby, B., Clements, B., Gupta, S., & Inchauste, G. (2004). The cyclical and long-term behavior of government expenditures in developing countries. *IMF working paper.*

Alesina, A., Perotti, R., Tavares, J., Obstfeld, M., & Eichengreen, B. (1998). The political economy of fiscal adjustments. *Brookings Papers on Economic Activity*, *1998(1)*, 197–266.

Alesina, A., & Perotti, R. (1995). The political economy of budget deficits. Staff Papers, 42(1), 1–31.

Alesina, A., Campante, F. R., & Tabellini, G. (2008). Why is fiscal policy often procyclical? *Journal of the European Economic Association, 6(5),* 1006–1036.

Alesina, A., Roubini, N., & Cohen, G. D. (1997). *Political cycles and the macroeconomy*. MIT press.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies, 58(2)*, 277–297.

Afonso, A., Agnello, L., & Furceri, D. (2010). Fiscal policy responsiveness, persistence, and discretion. *Public Choice, 145(3–4)*, 503–530.

Barro, R. J. (1979). On the determination of the public debt. *Journal of Political Economy, 87(5, Part 1)*, 940–971.

Bayar, A. H., & Smeets, B. (2009). Economic, political and institutional determinants of budget deficits in the European Union. *IMF working paper.*

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87(1)*, 115–143.

De Haan, J., & Mink, M. (2005). Has the stability and growth pact impeded political budget cycles in the European Union? *CESifo Working Paper Series.*

De Haan, J., & Sturm, J. E. (1997). Political and economic determinants of OECD budget deficits and government expenditures: A reinvestigation. *European Journal of Political Economy, 13(4)*, 739–750.

Franzese Jr, R. J. (2002). Electoral and partisan cycles in economic policies and outcomes. *Annual Review of Political Science, 5(1)*, 369–421.

Frankel, J. A., Vegh, C. A., & Vuletin, G. (2013). On graduation from fiscal procyclicality. *Journal of Development Economics, 100(1)*, 32–47.

Furceri, D., & Poplawski M. (2008). Fiscal volatility and the size of nations. *European Central Bank Working Paper*, 924.

Galí, J., & Perotti, R. (2003). Fiscal policy and monetary integration in Europe. *Economic Policy, 18(37)*, 533–572.

Henisz, W. J. (2004). Political institutions and policy volatility. *Economics & Politics, 16(1)*, 1–27.

Honda, J., & Miyamoto, H. (2020). Would Population Aging Change the Output Effects of Fiscal Policy? *International Monetary Fund.*

Lane, P. R. (2003). The cyclical behaviour of fiscal policy: evidence from the OECD. *Journal of Public Economics, 87(12)*, 2661–2675.

Leachman, L. L., Rosas, G., Lange, P., & Bester, A. (2007). The political economy of budget deficits. *Economics & Politics, 19(3)*, 369–420.

Makin, A. J., & Layton, A. (2021). The global fiscal response to COVID-19: Risks and repercussions. *Economic Analysis and Policy, 69*, 340–349.

Maltritz, D., & Wüste, S. (2015). Determinants of budget deficits in Europe: The role and relations of fiscal rules, fiscal councils, creative accounting and the Euro. *Economic Modelling, 48*, 222–236.

Mulas-Granados, C. (2003). The political and economic determinants of budgetary consolidation in Europe. *European Political Economy Review, 1(1)*, 015–039.

Perotti, R., & Kontopoulos, Y. (2002). Fragmented fiscal policy. *Journal of Public Economics, 86(2)*, 191–222.

Roubini, N., & Sachs, J. D. (1989). Political and economic determinants of budget deficits in the industrial democracies. *European Economic Review, 33(5)*, 903–933.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal, 9(1)*, 86–136.

Stojkoski, V., Utkovski, Z., Jolakoski, P., Tevdovski, D., & Kocarev, L. (2020). The socio-economic determinants of the coronavirus disease (COVID-19) pandemic. *arXiv preprint arXiv:2004.07947*.

Stojkoski, V., Jolakoski, P., & Ivanovski, I. (2020). The short-run impact of COVID-19 on activity in the insurance industry in the Republic of North Macedonia. *arXiv preprint arXiv:2011.10826*.

Talvi, E., & Vegh, C. A. (2005). Tax base variability and procyclical fiscal policy in developing countries. *Journal of Development Economics, 78(1)*, 156–190.

Tevdovski, D., Jolakoski, P., & Stojkoski, V. (2021). The impact of state capacity on the cross-country variations in COVID-19 vaccination rates. *Available at: SSRN 3801205*.

Volkerink, B., & De Haan, J. (2001). Fragmented government effects on fiscal policy: New evidence. *Public Choice, 109(3)*, 221–242.

Received: June 18, 2021
Accepted: January 15, 2022

APPENDIX

Table A1: List of variables and data sources

Variable	Source
Primary balance	IMF, WEO; General government primary net lending/borrowing % of GDP
Debt	IMF, WEO; General government gross debt % of GDP
Unemployment rate	IMF, WEO; % of total labour force
GDP	IMF, WEO; Gross domestic product at constant prices
Population	IMF, WEO in raw numbers
Government expenditure	IMF, WEO; % of GDP
Size of cabinet	CNTS (polit10)
Type of regime	CNTS (polit02)

Table A2: Results without the interest rate: Blundell–Bond linear dynamic panel-data estimation

	(1)	(2)	(3)	(4)
VARIABLE	Macroeconomic	Demographic	Political/Institutional	2020
Primary balance (t-1)	0.591***	0.570***	0.550***	0.578***
	(0.025)	(0.031)	(0.058)	(0.041)
GDP growth rate	0.398***	0.371***	0.366***	0.213***
	(0.018)	(0.019)	(0.019)	(0.018)
Unemployment rate	0.001	–0.008	–0.076	–0.170***
	(0.062)	(0.056)	(0.086)	(0.064)
Government debt	1.212	1.046	1.736	3.466*
	(0.774)	(0.742)	(1.196)	(1.773)
Government expenditure	–7.595***	–10.035***	–10.189***	–8.294***
	(1.213)	(1.380)	(1.302)	(1.988)
Population		–3.275***	–1.707*	–2.776*
		(0.732)	(1.036)	(1.531)
Polity			0.492*	0.583*
			(0.295)	(0.342)
Size of cabinet			–0.049***	–0.030**
			(0.015)	(0.015)

Type of regime (2)			2.876	7.737**
			(4.128)	(3.508)
Type of regime (3)			6.080	9.799
			(9.925)	(10.018)
GDP*2020				–0.209
				(0.296)
Unemployment*2020				0.551***
				(0.161)
Government expenditure*2020				–2.387*
				(1.358)
Debt*2020				0.180
				(1.686)
Population*2020				0.218
				(0.504)
Constant	21.170***	39.878***	30.732***	19.763*
	(4.319)	(8.096)	(7.065)	(10.754)
Observations	1,002	1,002	1,002	1,002
Number of countries	43	43	43	43
Arellano-Bond test	p = 0.747	p = 0.705	p = 0.582	p = 0.736
Sargan test (H0: over-identifying restrictions are valid)	p = 1.000	p = 1.000	p = 1.000	p = 1.000

Note: Standard errors in parentheses

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

Table A3: List of countries and their fiscal measures, primary budget balance, and gross central government debt in domestic currency denomination, as a percentage of GDP

Country	Fiscal measures	Primary balance	Debt in domestic currency
Albania	2.87	–8.42	33.53
Armenia	2.98	–5.82	10.90
Australia	13.53	–10.06	51.55
Brazil	14.61	–16.78	69.90
Bulgaria	6.66	–2.00	4.20
Canada	16.73	–19.92	42.69
Colombia	5.27	–9.48	32.79
France	20.98	–10.77	95.47
Hungary	8.55	–8.28	56.17
Indonesia	3.83	–6.32	18.92
Ireland	7.59	–6.00	65.48
Israel	9.67	–12.94	52.12
Italy	37.93	–12.98	133.14
Lithuania	8.28	–6.72	35.78
Luxembourg	11.16	–6.98	20.89
Mexico	1.10	–5.8	28.97
Moldova	2.63	–8.00	11.21
Netherlands	8.85	–8.76	51.10
Philippines	3.37	–8.06	0.63
Portugal	9.95	–8.35	126.24
Romania	5.38	–9.59	19.51
Russia	3.41	–5.29	9.36
Slovak Republic	6.97	–8.84	51.82
Slovenia	14.56	–8.82	68.01
South Africa	9.61	–14.04	55.45
Spain	17.66	–14.09	92.06
Sweden	10.63	–5.90	31.48
Thailand	12.46	–5.21	35.58
Turkey	13.75	–7.88	21.10
United Kingdom	25.72	–16.46	86.99
United States	14.22	–18.72	101.88

ECONOMIC ANNALS, Volume LXVII, No. 232 / January – March 2022
UDC: 3.33 ISSN: 0013-3264

https://doi.org/10.2298/EKA2232127A

*Adedeji Adeniran**
*Mma Amara Ekeruche***
*Chimere O. Iheonu****

THE QUALITY OF BUDGETARY INSTITUTIONS IN AFRICA: EXPLORING THE DRIVERS

ABSTRACT: *This paper examines the economic and political determinants of the observed variations in the quality of budgetary institutions in 31 selected African countries from 2005 to 2017. The quality of budgetary institutions is measured using the World Bank's Country Policy and Institution Assessment score. The empirical analysis utilises Ordinary Least Squares, two-stage least squares, two-step generalized method of moment, and the random effects probit and mixed effects models. The most significant and robust determinants of budgetary institution quality were found to be the level of external debt, foreign aid, the extent of control of corruption, and the level of voice and accountability. The results also reveal that foreign aid, control of corruption, and voice/accountability increase the probability of an improved quality of budgetary institutions. These findings resonate with the broader discussion on the role of political will and the need for the preferences of domestic actors to be aligned in order to deliver institutional reform in Africa.*

KEY WORDS: *budgetary institution, crisis hypothesis, common pool problem, fiscal performance, Africa*

JEL CLASSIFICATION: C23, H61, P16

* Centre for the Study of the Economies of Africa, Abuja, Nigeria, email: aadedeji@cseaafrica.org
** Centre for the Study of the Economies of Africa, Abuja, Nigeria, email: mma.ekeruche@cseaafrica.org
*** Centre for the Study of the Economies of Africa, Abuja, Nigeria, email: ciheonu@cseaafrica.org

1. INTRODUCTION

Budgetary institutions play a fundamental role in a country's economic development and prosperity. First, the efficient and effective utilisation of public resources, the strategic alignment of finances with a country's priorities, and the maintenance of a sustainable fiscal programme all require sound budgetary institutions. Second, fiscal management and stabilisation policies, especially the adoption of counter-cyclical fiscal policies, are dependent on the quality of budgetary institutions. Third, budgetary institutions are also critical for ensuring government accountability and transparency and delivering pro-poor policies. According to Raudla (2014), budgetary institutions cover two aspects of institutional arrangements. The first is fiscal rules, which entail constraints in taxation, debt, and public spending, and the second is budgetary process rules, the procedures associated with public budgeting which outline the process that governs the preparation, adoption, and implementation of the budget.

Several instances demonstrate the importance of high-quality budgetary institutions to key economic and development indicators. Gollwitzer (2010) reveals that budgetary institutional quality is associated with better fiscal performance – specifically, low primary budget balance. In Latin America, countries whose budgetary institutions rank within the top 25% (with regards to rigidity of fiscal rules and hierarchical and transparent procedures) achieve a better fiscal balance of about 2 percentage points of GDP than countries within the bottom 25% (File & Scanticini, 2004). Similarly, using a global sample, Piatti-Fünfkirchen and Smets (2019) conclude that improved public finance management and, by extension, strengthening of budgetary institutions, reduce under-5 mortality by 14 deaths per 1,000 live births.

Despite the broad consensus on the importance of budgetary institutions for economic performance, there is limited understanding of what makes budgetary institutions more viable in some countries than in others. The limited evidence on the determinants of the quality of budgetary institutions mostly refers to developed countries (Von Hagen, 2002 & Mulas-Granados, 2003 for Europe) or inferences that can be drawn from the larger literature on political institutions. Despite the remarkable heterogeneity in the quality of budgetary institutions in Africa (see Gollwitzer, 2010), attempts to investigate the economic and political factors behind this variation are limited. While the pivotal role of budgetary

institutions in fiscal sustainability in Africa is widely established (see Dabla-Norris, Allen, Zanna, Prakash, Kvintradze, Lledo, Yackovlev & Gollwitzer, 2010), studies of the causal factors behind quality budgetary institutions are scarce and the few that do exist rely on qualitative analysis of budgetary institutions in individual countries (see Department for International Development (DFID), 2014 for Nigeria).

This paper fills this gap by empirically investigating the economic and political drivers of the observed variation in the quality of budgetary institutions in Africa. Following the World Bank (2019), we define high-quality budgetary institutions as those characterized by (1) adoption of rules that specify transparent budgeting processes, (2) allowing for design and implementation of effective fiscal policies, and (3) allowing for fiscal reporting in a comprehensive and timely fashion. The World Bank's Country Policy and Institution Assessment (CPIA-13) codifies these characteristics into a measurable indicator by ranking budgetary institutions across countries on a 6-point scale. Using this indicator for selected African countries between 2005 and 2017, we find that around 51% of these countries scored above 3.2, the benchmark for high-quality budgetary institutions. This means that across the continent there are an equal number of high-quality and weak budgetary institutions.

We find that high-quality budgetary institutions are more likely to be present in countries with low external debt levels, higher levels of foreign aid, lower levels of corruption, and better accountability levels. We also find strong evidence that foreign aid, the control of corruption, and voice/accountability increase the probability of improving the quality of budgetary institutions in Africa. Overall, political factors have a more robust effect on budgetary institutions than economic factors. This concurs with the politico-institutional hypothesis that budgetary institutions are political instruments (Alesina & Perotti, 1995). The implication of this is that since the identified political factors are within the domain of domestic state actors, improving the quality of budgetary institutions depends largely on political will.

The paper contributes to the literature in two notable areas. First, we are not aware of any previous quantitative study on the drivers of the quality of budgetary institutions in the African context. Given the recurring debt problem experienced

in many African countries over the years, understanding the policy menu that is available for strengthening budgetary institutions can result in better fiscal management. Second, this study relates to the emerging literature on the importance of institutions for economic development. For example, Acemoglu and Robinson (2012) argue that the creation of inclusive economic and political institutions is an offshoot of economic development. Given that budgetary institutions are central to this, the findings from this study will deepen the understanding of how institutions work and affect the economy.

The remainder of the paper is structured as follows. Section 2 reviews the literature on the factors that are associated with the quality of budgetary institutions. Details on the data sources and the methodology are presented in section 3. The findings of the study are discussed in section 4, while section 5 provides concluding remarks.

2. LITERATURE REVIEW

The literature has identified several factors that are associated with the quality of institutions in general and budgetary institutions in particular. One of the well-established linkages in this regard is between the quality of budgetary institutions and the extent of the common pool problem in fiscal policymaking (Weingast, Shepsle, and Johansen, 1981; Velasco, 2000). The public budget is a common pool belonging to many political constituencies and this institutional structure leads to the different constituencies competing for budgetary resources without internalising the costs involved. This creates spending and deficit biases, thereby causing fiscal crisis. The implication is that weak budgetary institutions prevail in countries that are fractionalised or lacking checks and balances in the form of strong electoral competition, because the common pool problem is greater in these settings (Von Hagen, 2002). Similarly, the political structure that mitigates the common pool problem will positively influence the quality of budgetary institutions. De Renzio and Cho (2020) show that fiscal transparency and timely and comprehensive reporting are critical for budget credibility, and that political and economic factors such as democracy and income do not influence the credibility of budgets. Further, Heinsz (2004) suggests that the presence of checks and balances, which the parliamentary system offers, may improve economic outcomes.

Another body of work with profound implications for the trajectory of budgetary institutions is the literature on the resource curse. Specifically, Sachs and Warner (2001) found that countries with abundant natural resources are more likely to have weak institutions and procyclical fiscal policy. This suggests that poor quality budgetary institutions will be prevalent in countries that are resource dependent. However, in a few resource-endowed countries in the developing world the presence of natural resources has promoted growth and not weakened their institutions. Acemoglu, Johnson, and Robinson (2003) showed that in the previous 35 years, Botswana, with its diamonds, had the highest per capita growth globally, which Cabrales and Hauk (2011) link to the presence of good quality institutions before the emergence of natural resources. Other manifestations of poor institutional quality such as lack of transparency, corruption, and limited voice and accountability for citizens will affect budgetary institutions directly and indirectly. Schick (1998) noted that corruption is a crucial factor contributing to the poor budget management observed in Africa and other low-income countries. Similarly, Alesina and Tabellini (2008) find that in developing countries the majority of procyclical fiscal policies – a measure of the strength of budgetary institutions – can be explained by high levels of corruption. Furthermore, Alt and Lassen (2006) find that higher transparency in the political process reduces state actors' ability to be corrupt. Ngo and Nguyen (2020) reveal the need to improve the institutional set-up of economies to avoid high and unstable budget deficits. The tendency for corruption and absence of voice and accountability to reduce the quality of budgetary institutions reinforces the agency problem examined in Persson and Tabellini (2000), where politicians appropriate resources for personal gain at the expense of citizens.

Furthermore, the literature on the 'crisis hypothesis' suggests that the likelihood for reform is higher during economic crises (Mahmalat, & Curran, 2018). For instance, recession provides an opportunity to introduce structural reforms like budget transparency or fiscal rules, which are unlikely during an economic boom. Therefore, growth trends should influence the quality of budgetary institutions. Gradstein (2008) supports this hypothesis, as he finds that backward economic development, measured by income per capita, is associated with poor institutional quality. Similarly, the level of debt can also induce public management and fiscal reforms to avert debt crisis. Ayee (2008) argues that most administrative and public financial management reforms in Africa result from

serious economic crises. This is at least the case with the introduction of the medium-term fiscal framework and fiscal rules in Uganda, Nigeria, and Ghana in response to debt crisis (World Bank, 2013; Akunyili, Katz & Duncan, 2013).

The role of trade openness in institutional quality is inconclusive. Rodrik (2000) argues that the adoption of trade liberalisation policies, which in many cases entails the adoption of specific institutional norms, assists in improving domestic institutions, including budgetary institutions. Similarly, Islam and Montenegro (2002) find that trade openness produces better institutions because rent-seeking and corruption are more difficult when there is increased competition between agents. Do and Levchenko (2009) observe that international trade could contribute to the concentration of power in the hands of individuals that intend to establish or perpetuate bad institutions. The influence of foreign aid inflow on budgetary institutions is also mixed. On the positive side, foreign aid can introduce some form of accountability into the budgeting process in response to aid conditionalities from donors. Aid components going into institutional/technical support will have a direct positive impact on budgetary institutions. For example, Nigeria was able to establish a Budget Office, Debt Management Office, and other fiscal reforms largely due to donor support (Akunyili, Katz & Duncan, 2013). Similarly, in a study of 53 African countries, Asongu (2015) finds that there are institutional benefits from foreign aid inflows; however, these benefits are contingent on existing institutional levels. A few earlier studies including Easterly (2003) and Islam (2003) also argue that foreign aid improves institutional quality by allowing the government to invest in activities that improve the quality of bureaucracy, reduce corruption, and enforce the rule of law. However, there are concerns that aid dependency can weaken institutions. Examining the relationship between aid and institutions in sub-Saharan Africa (SSA), Brautigam and Knack (2004) argue that large amounts of aid given over long time periods can weaken institutions.

The approach of this study is to investigate both economic and political factors that influence the quality of budgetary institutions using various econometric procedures for robustness purpose. There are very few studies on the subject matter regarding Africa, yet weak budgetary institutions are more prevalent on that continent than in other regions (Dabla-Norris et al., 2010). These identified gaps are the motivation for this study.

3. DATA AND METHODOLOGY

3.1 Measurement of budgetary institution quality

Several indicators to measure the quality of budgetary institutions have been proposed in the literature. These include, but are not limited to, Public Expenditure and Financial Accountability (PEFA) indicators, the Open Budget Index by the International Budget Partnership, and the Quality of Budgetary and Financial Management criteria (CPIA-13) – a sub-index of the World Bank's Country Policy and Institution Assessment (CPIA) indicators. This paper employs CPIA-13 as our measure of the quality of budgetary institutions. CPIA-13 is an expert scoring of a country's budgetary system on a scale of 1 (worst performance) to 6 (best performance). The CPIA-13 is a composite index made up of three components: (1) comprehensiveness and credibility of the budget and its link to policy priorities, (2) effectiveness of financial management systems to ensure that the budget is implemented in a predictable manner, and (3) timeliness and accuracy of accounting and fiscal reporting, including timely audit of public accounts. The key advantage of CPIA-13 is that it has wider coverage across both time and country than other indicators of budgetary institutions. This ensures that we have enough data points for empirical analysis. However, the CPIA-13 score is not disaggregated along its key components, which means our analysis is limited to identifying the effect of specific economic and political factors on only the overall quality of budgetary institutions.

Given that the CPIA-13 scores are ranked, to improve the robustness of the results it is important for empirical analysis to specify the benchmark in order to separate the countries with high-quality budgetary institutions from the others. Following IMF (2009), which stipulates the mean values of CPIA scores, we set the threshold for a high-quality budgetary institution at 3.2 and above for countries under the World Bank's Poverty Reduction and Growth Facility, and 3.1 and above for countries under the International Development Association (IDA). We called this approach threshold definition. The study also makes use of the actual score in CPIA-13 – a cluster-level analysis. To further test the robustness of the results we also employ a quintile definition, where quality of budgetary institutions is considered high if a country is above the bottom two quintiles of CPIA-13 scores. Quintile definition is widely used in benchmarking

the CPIA score, especially for computation and classification of state fragility (see Bertocchi & Guerzoni, 2010).

Table 1 summarizes the CPIA-13 score between 2005 and 2017 for the 31 SSA countries for which the data is publicly available. The average score for SSA is 3.19, but the score varies widely with a recorded minimum of 1.5 and maximum of 4.5. No country has reached the maximum score of 6, which means there is still enormous room for improvement even in countries deemed to have high-quality budgetary institutions. According to the threshold definition of quality, 203 of the 399 budgetary institutions (50.8%) are high-quality, and this increases to 75.9% of the observations if quintile definition is used instead. This means that a significant number of countries within the bottom two quintiles score below the 3.2 threshold. In essence, the threshold definition sets a higher benchmark for measuring the quality of budgetary institutions than quintile definition.

Table 1: Summary statistics for CPIA-13

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Overall CPIA	399	3.1867	0.6022	1.5	4.5
Threshold definition					
High quality	203	3.6822	0.2748	3.5	4.5
Low quality	196	2.6734	0.3762	1.5	3
Quintile definition					
High quality	303	3.6822	0.2748	3.5	4.5
Low quality	96	2.3333	0.2477	1.5	2.5

Source: Authors computation

3.2 Methodology and Model specification

We model the determinants of the quality of budgetary institutions as a function of economic and political factors using traditional panel data econometric specification as follows:

$$qbi_{i,t} = \alpha_0 + \beta E_{i,t} + \varphi P_{i,t} + \varepsilon_{i,t} \quad (1)$$

where $qbi_{i,t}$ represents the measure of the quality of budgetary institutions for country *i* in period *t*. qbi is measured in two ways: the actual CPIA-13 score is

first used for cluster analysis and then as a binary dummy variable which takes a value of one for high-quality budgetary institution and zero otherwise in both threshold and quintile definitions. Different measures of budgetary institutions require different estimation approaches; hence, a battery of econometric techniques will be used to estimate Equation (1). $E_{i,t}$ represents the vector of economic factors, which includes the GDP growth rate, trade openness as measured by imports plus exports as a percentage of GDP, external debt stock, total GDP to account for economic size, aid as a percentage of GDP, and natural resource rent. $P_{i,t}$ represents the vector of political factors, including a fractionalisation variable to capture the degree of ethnic, linguistic, and religion fragmentation, level of political competition, system of government (parliamentary or presidential), level of corruption control, and voice and accountability. The link between the economic and political factors utilised in this study and the quality of budgetary institutions has been documented in the literature review section. $\varepsilon_{i,t}$ represents the standard error term. Table 2 contains the summary statistics for and definition of all the variables.

Table 2: Descriptive statistics

Variable	Description	Mean	Std. Dev.	Min	Max
GDP growth	Real GDP growth	5.0888	4.4756	–36.6999	20.7157
Trade openness	Import plus export as % of GDP	67.9143	28.0595	19.10	165.6459
External debt stock	Total external debt to GDP	33.5333	26.9949	0	268.4864
Size of the economy	Real GDP constant United States dollar	23.1016	1.4675	18.64449	27.0662
Aid	Development assistance and official aid received (current US$) as % of GDP	0.0891	0.0632	0.0016	0.3587
Natural resource rents	Total natural resource rents (% of GDP)	13.4643	11.4599	0.4500	60.1212
Fractionalisation	Average of ethnic, linguistic, and religious fractionalisation, taken from Alesina and Ferrara (2005)	0.6491	0.1688	0.2094	0.8404
System of governance	0=Presidential; 1=Parliamentary	0.0645	0.2459	0	1
Political competition	Share of the vote of the ruling party (the variable takes 1 for non-democratic states)	0.6989	0.1702	0.2620	1
Control of corruption	Country score on World Bank's Corruption Index	–0.6933	0.5408	–1.5447	0.9495
Voice and accountability	Country score on World Bank's Voice & Accountability Index	–0.5633	0.6525	–1.8301	0.9861

Source: Authors computation

Finally, we address a potential issue that could affect the validity of our results. The possibility of an endogeneity problem is high in our model due to reverse causality/simultaneity between the quality of budgetary institutions and the explanatory variables. For example, the quality of budgetary institutions is a major determinant of debt level and economic growth. To account for this, the study utilised the two-stage least squares (2SLS) and the two-step generalized method of moments (GMM) methodologies to account for endogeneity, along with the traditional Ordinary Least Squares estimator which is heteroskedastic and autocorrelation-consistent (HAC) in standard errors. The 2SLS and GMM estimation procedures utilise internal instruments using lags of the explanatory variables. According to Hayashi (2000) as revealed by Okwoche and Iheonu (2021), the two-step GMM has relative efficiency gains over the traditional 2SLS method. However, dummy explanatory variables such as system of governance and political competition are not instrumented. HAC standard errors are also employed in the 2SLS and the GMM specifications. Prior to estimating the regression model, the study tests for multicollinearity using a simple correlation matrix and cross-sectional dependence using the Pesaran (2015) procedure. According to Baltagi, Kao, and Peng (2016), cross sectional dependence entails common shocks across sectional units such that the error term in country *i* is correlated with that of country *j*. According to Nathaniel and Iheonu (2019), not accounting for cross-sectional dependence in an econometric model can lead to estimation bias. This study utilises the Driscoll and Kraay (1998) regression which accounts for cross-sectional dependence as a robustness strategy.

The study also employs the Random Effects (RE) probit model and the Mixed Effects (ME) probit model to analyse the probability of improving or reducing the quality of budgetary institutions in Africa, which is dependent on the regressors in the models. According to Ekeocha and Iheonu (2020), the RE probit model in this study context is such that a standard normal distribution function for the nonlinear parameterisation at their means is imposed to ascertain the predicted probability of improving the quality of budgetary institutions in Africa. The ME probit model, however, side-steps the assumption of dealing with time-invariant effects using the Fixed Effects or RE model and also yields robust estimates when analysing categorical data. Due to the nature of the data, which is unbalanced with a large number of cross sections and fewer time periods, the stationarity test is not a relevant statistical analysis. This is because the test for

stationarity is more consistent in macro panel data where the number of time periods is significantly large. This is consistent with extant literature where the number of cross sections is larger than the number of time periods (Iheonu & Ichoku, 2021; Asongu & Odhiambo, 2020; Asongu, Nnanna & Acha-Anyi, 2020).

4. PRESENTATION AND DISCUSSION OF RESULTS

This section begins with the correlation matrix of the variables utilised in our econometric models. The correlation matrix is evaluated as a simple measure of understanding whether multicollinearity exists in the regressions. The findings are such that we do not see high levels of correlation between the explanatory variables utilised in the models, except for the two indicators of institutional quality: the correlation between voice/accountability and the control of corruption is 0.72. This means that for unbiased estimates to be derived in the regressions, these variables have to be in separate models.

Table 3: Correlation Matrix of the Variables

	CPIA	CPIA 1	CPIA 2	GDP	Trade	Debt	Eco Size	Aid	Resource Rent	Frac. Index	Gov. System	Pol. Comp.	VA	CC
CPIA	1.00													
CPIA 1	0.84	1.00												
CPIA 2	0.82	0.65	1.00											
GDP	0.19	0.15	0.14	1.00										
Trade	−0.24	−0.24	−0.26	−0.01	1.00									
Debt	−0.33	−0.23	−0.36	−0.23	0.29	1.00								
Eco Size	−0.12	−0.16	−0.11	0.15	−0.17	−0.38	1.00							
Aid	0.42	0.37	0.38	0.08	−0.17	0.17	−0.61	1.00						
Resource Rent	−0.39	−0.41	−0.46	0.17	0.40	−0.02	0.35	−0.31	1.00					
Frac. Index	−0.17	−0.17	−0.07	0.03	−0.27	−0.13	0.41	−0.24	0.25	1.00				
Gov. System	0.05	0.04	0.12	0.05	0.38	−0.01	−0.13	−0.01	−0.10	−0.22	1.00			
Pol. Comp.	−0.01	0.02	−0.05	−0.20	−0.18	−0.05	0.17	−0.19	−0.15	−0.03	−0.07	1.00		
VA	0.66	0.63	0.67	0.07	0.02	−0.20	−0.18	0.25	−0.43	−0.22	0.07	−0.25	1.00	
CC	0.65	0.56	0.58	0.07	0.07	−0.12	−0.38	0.40	−0.50	−0.59	0.23	−0.15	0.72	1.00

Source: Authors computation.

Note: CPIA 1 is the Threshold Definition, CPIA 2 is the Quintile Definition, Eco Size is Economy Size, Frac. Index is Fractionalisation Index, Gov. System is System of Government, Pol. Comp. is Political Competition, VA is Voice and Accountability, CC is Control of Corruption.

The indicators of the quality of budgetary institutions are positively correlated with GDP, foreign aid, system of governance, voice and accountability, and control of corruption. However, the indicators of the quality of budgetary institutions have a negative correlation with trade openness, external debt, size of the economy, resource rent, and the fractionalisation index.

Table 4 shows the study test for cross-sectional dependence in the econometric model, utilising the Pesaran (2015) cross-sectional dependence test. Cross-sectional dependence is present in the model at the 10% level of statistical significance for control of corruption and at the 5% level of statistical significance for voice/accountability.

Table 4: Pesaran (2015) Cross-sectional Dependence Test

Model	Test Value	Probability
Model with Control of Corruption	1.9720*	0.07
Model with Voice and Accountability	2.0110**	0.04

Source: Authors computation.

Note: ** and * represent statistical significance at 5% and 10%.

The findings of this study are presented in Table 5 and cut across various econometric methods. Additional findings are revealed in the Driscoll and Kraay regression, which accounts for cross-sectional dependence and also HAC in standard errors. We find that the GDP growth rate increases the quality of budgetary institutions in Africa. However, the significance of the GDP growth rate differs across the methods utilised. While GDP growth rate significantly increases budget quality in Africa, accounting for endogeneity using the 2SLS and the GMM methodologies reveals that GDP growth rate does not significantly influence budget quality. The positive relationship between GDP growth rate and the quality of budgetary institutions is due to the fact that increased GDP increases fiscal space, which in turn leads to lower levels of debt and fiscal crisis. Furthermore, trade openness is revealed to have a negative influence on the quality of budgetary institutions. However, the negative impact is insignificant when endogeneity is accounted for. The study finds that external debt significantly reduces the quality of budgetary institutions in Africa. The negative impact of external debt on budgetary institutions is in line with the common pool problem due to the creation of increased spending and deficit biases.

Additionally, foreign aid is positively and significantly associated with improved budgetary institutions. One explanation as to why countries that receive large inflows of aid have stronger budgetary institutions is that donor countries as well as multilateral development banks (MDBs) have focused on strengthening the budgetary institutions in recipient countries in the past two decades (Dabla-Norris et al., 2010). The result is consistent with the study of Asongu (2015). The findings also show that political competition, control of corruption, and voice/accountability significantly improve the quality of budgetary institutions, even after endogeneity has been considered. The finding on the corruption and budgetary institution nexus is in line with the findings of Schick (1998) and Alesina and Tabellini (2008). Alt and Lassen (2006) and Ngo and Nguyen (2020) have also revealed the importance of institutional quality in improving the quality of budgetary institutions. Improving institutional quality reduces bottlenecks and increases efficiency, which enhances the quality of budgetary institutions.

The validity of the instruments in the 2SLS and GMM regressions are evaluated using the Sargan test. Findings reveal that the instruments utilised are valid, as the probability values are greater than the conventional levels of statistical significance. The Kleibergen-Paap rk LM statistic for weak instruments reveals strong evidence that the 2SLS and the GMM models are identified.

Table 5: Panel Data Regressions

Variable	OLS	OLS	2SLS	2SLS	GMM	GMM
GDP	0.0119*	0.0145**	0.0212	0.0451	0.0230	0.0039
	(0.087)	(0.029)	(0.564)	(0.238)	(0.524)	(0.891)
Trade	–0.0021	–	–0.0023	–0.0051	–0.0022	–0.0033
	(0.235)	0.0042**	(0.526)	(0.196)	(0.525)	(0.373)
		(0.015)				
External Debt	–	–	–	–	–	–
	0.0055**	0.0055**	0.0096**	0.0090**	0.0096**	0.0106***
	(0.012)	(0.011)	(0.010)	(0.023)	(0.010)	(0.004)
Economy Size	0.0606	0.0327	0.0691	0.0356	0.0669	0.0815
	(0.120)	(0.379)	(0.290)	(0.600)	(0.302)	(0.197)
Aid	3.4243***	4.0894***	4.4151***	5.6279***	4.3641***	6.4654***
	(0.000)	(0.000)	(0.009)	(0.001)	(0.009)	(0.000)
Resource Rent	0.0005	0.0033	0.0021	0.0073	0.0021	0.0047
	(0.887)	(0.401)	(0.751)	(0.333)	(0.753)	(0.508)
Fractionalisation	1.0463***	–0.1953	0.7612**	–0.5381*	0.7650**	–0.5753*
	(0.001)	(0.445)	(0.041)	(0.091)	(0.040)	(0.063)
System of Governance	–0.0573	0.2877**	0.0290	0.4363	0.0295	0.2894
	(0.579)	(0.012)	(0.937)	(0.280)	(0.936)	(0.450)
Political Competition	0.5076**	0.7190***	0.5743	0.9278**	0.5874*	0.6031*
	(0.028)	(0.002)	(0.110)	(0.020)	(0.099)	(0.076)
Control of Corruption	0.9307***		0.9396***		0.9401***	
	(0.000)		(0.000)		(0.000)	
Voice and		0.6336***		0.6499***		0.6151***
Accountability		(0.000)		(0.000)		(0.000)
Constant	1.3166	2.3727**	1.2113	2.1860	1.2452	1.4693
	(0.180	(0.010)	(0.423)	(0.165)	(0.409)	(0.333)
Centred R^2	0.6360	0.6413	0.6525	0.6150	0.6511	0.6503
Uncentered R^2	0.9848	0.9850	0.9850	0.9834	0.9850	0.9849
F-statistic	32.95***	27.60***	21.52***	21.08***	21.55***	21.99***
	(0.0000)	(0.000)	(0.000)	(0.000)	(0.0000)	(0.0000)
Kleibergen-Paap (KP)			9.547***	9.468***	9.547***	13.990***
LM statistic			(0.0085)	(0.0088)	(0.0085)	(0.0029)
Sargan statistic			0.103	0.210	0.103	3.348
			(0.7479)	(0.6471)	(0.7479)	(0.1875)
Observations	226	226	171	171	171	171

Source: Authors computation. **Note:** ***, **, and * denote statistical significance at 1%, 5%, and 10% respectively. Results are HAC in standard errors.

Table 6: Driscoll and Kraay Regression

Variable	Driscoll and Kraay (1)	Driscoll and Kraay (2)
GDP	0.0291*	0.0380***
	(0.095)	(0.001)
Trade	–0.0019	–0.0033***
	(0.115)	(0.000)
External Debt	–0.0072***	–0.0072***
	(0.003)	(0.000)
Economy Size	0.0640***	0.0470***
	(0.001)	(0.003)
Aid	4.8730***	5.7007***
	(0.000)	(0.000)
Resource Rent	–0.0002	0.0010
	(0.940)	(0.564)
Fractionalisation	1.0292***	–0.1751
	(0.000)	(0.338)
System of Governance	0.0145	0.3121***
	(0.866)	(0.000)
Political Competition	0.5044***	0.6903***
	(0.000)	(0.000)
Control of Corruption	0.9259***	
	(0.000)	
Voice and Accountability		0.6147***
		(0.000)
Constant	1.0555**	1.7747***
	(0.034)	(0.002)
R^2	0.6711	0.6731
F-statistic	25145.66***	752.75***
	(0.0000)	(0.0000)
Observations	199	199

Source: Authors computation. **Note:** ***, **, and * represents statistical significance at 1%, 5%, and 10% respectively. Probability values are in parenthesis.

In Table 6, cross-sectional dependence is accounted for using the Driscoll and Kraay (1998) regression. The regressors in the models are instrumented using their first lags to account for endogeneity, except for the system of governance

and political competition variables. The findings are similar to those in the previous table, except that this time we have a larger number of significant variables. Consistent with the results in Table 5, external debt, foreign aid, control of corruption, and voice/accountability all significantly influence the quality of budgetary institutions in Africa. We also find that the influence of political competition on budget quality is sensitive to the measure of institutional quality utilised.

The study uses the marginal effect of the Probit RE and ME models and the threshold approach to the CPIA to capture the probability that the quality of budgetary institutions in Africa has improved. The first lags of the regressors in the models are used as instruments in the original regression to account for possible endogeneity. However, system of governance and political competition are not instrumented due to the binary nature of the variables. Table 7 shows that trade openness reduces the probability that the quality of budgetary institutions in Africa has improved. The results also show that control of corruption, voice/accountability, and foreign aid increase the probability of improved budgetary institutions in Africa. However, the study did not find that external debt, economy size, or resource rent had any significant influence on the probability of improving budgetary institution quality. African countries with a parliamentary system of governance have a higher probability of improving budgetary institutions than those with a presidential system. This finding is in line with Heinsz (2004). This is significantly true when voice/accountability is the indicator of institutional quality rather than control of corruption.

Table 7: Marginal Effect of Probit Regressions (Threshold Approach) CPIA >3.2 or 3.1

Variable	Probit RE Model	Probit RE Model	Probit ME Model	Probit ME Model
GDP	–0.0045	–0.0087	0.0041	0.0048
	(0.584)	(0.254)	(0.683)	(0.571)
Trade	–0.0053***	–0.0061***	–0.0036**	–0.0042***
	(0.005)	(0.001)	(0.040)	(0.007)
External Debt	–0.0005	–0.0015	–0.0012	–0.0013
	(0.768)	(0.312)	(0.437)	(0.389)
Economy Size	–0.0068	–0.0192	0.0154	–0.0182
	(0.895)	(0.693)	(0.640)	(0.543)
Aid	0.1156	0.1599**	0.1967***	0.1690***
	(0.155)	(0.025)	(0.007)	(0.006)
Resource Rent	–0.0005	0.0022	–0.0035	0.0028
	(0.936)	(0.741)	(0.477)	(0.498)
Fractionalisation	0.3161	–0.2707	0.6236**	–0.0694
	(0.529)	(0.469)	(0.025)	(0.737)
System of	0.1840	0.4017*	0.1947	0.3273**
Governance	(0.478)	(0.055)	(0.343)	(0.041)
Political	0.0447	0.1517	0.3951**	0.4744***
Competition	(0.829)	(0.456)	(0.029)	(0.002)
Control of	0.5495***		0.5639***	
Corruption	(0.000)		(0.000)	
Voice and		0.3589***		0.3846***
Accountability		(0.000)		(0.000)
Number of	199	199	199	199
Observations				
Log likelihood	–44.7054	–44.8854	–73.2284	–62.4441

Source: Authors computation. **Note:** ***, **, and * denote statistical significance at 1%, 5%, and 10% respectively. Probability values are in parenthesis.

Table 8 shows the marginal effect results using the quintile definition of the CPIA score. This analysis is used as a robustness check for the threshold approach. The signs of the coefficient across the models and estimation technique are similar.

The results show that trade openness, external debt, and resource rent have a negative influence on the quality of budgetary institutions in Africa.

Table 8: Marginal Effect of Probit Regressions (Quintile Approach) CPIA>3

Variable	RE Probit Model	RE Probit Model	ME Probit Model	ME Probit Model
GDP	0.0068	0.0012	0.0117*	0.0011
	(0.343)	(0.799)	(0.056)	(0.792)
Trade	–0.0004	–0.0008	–0.0004	–0.0008
	(0.701)	(0.397)	(0.593)	(0.257)
External Debt	–0.0011	–0.0018	–0.0013	–0.0018**
	(0.282)	(0.301)	(0.133)	(0.016)
Economy Size	0.0488**	0.0209	0.0458***	0.0209*
	(0.044)	(0.239)	(0.003)	(0.089)
Aid	0.1703***	0.2008*	0.1926***	0.2008***
	(0.001)	(0.081)	(0.000)	(0.000)
Resource Rent	–0.0038	–0.0008	–0.0039	–0.0008
	(0.267)	(0.699)	(0.126)	(0.676)
Fractionalisation	0.8524**	0.3702	0.9077***	0.3702***
	(0.010)	(0.261)	(0.000)	(0.002)
System of Governance	–	–	–	–
Political Competition	0.0885	0.4749	0.2579*	0.4750***
	(0.690)	(0.304)	(0.065)	(0.000)
Control of Corruption	0.5450***		0.5502***	
	(0.000)		(0.000)	
Voice and Accountability		0.4086***		0.4087***
		(0.000)		(0.000)
Number of Observations	193	193	193	193
Log likelihood	–23.4578	–15.8165	–26.3008	–15.8165

Source: Authors computation. **Note:** ***, **, and * denote statistical significance at 1%, 5%, and 10% respectively. Probability values are in parenthesis.

It is further revealed that GDP growth rate, size of the economy, foreign aid, fractionalisation index, control of corruption, and voice/accountability all have a

positive influence on the quality of budgetary institutions in Africa. The result from the quintile approach is consistent with the findings of the threshold approach in terms of the relationship between trade openness, external debt, foreign aid, political competition, control of corruption, and voice/accountability. Low levels of corruption in general, and the design and implementation of the budget through budget transparency and opportunities for citizen participation, are likely to ensure that budgets are credible, and that proposed and actual spending align.

Comparing the results across all three measures of budgetary institutions (actual CPIA-13 score, the threshold definition, and the quintile definition), certain factors consistently play a role: the level of exposure to foreign aid, political competition, extent of control of corruption, and voice and accountability. This points to crucial policy windows for improving budgetary institutions in Africa. For instance, curbing corruption by streamlining the process of budget design, approval, and implementation, sanctioning the use of public funds for private gain, and encouraging open and continuous exchanges between key stakeholders will improve budget use. Additionally, the finding that development assistance to African countries is associated with improvements in budgetary institutions underscores a possible role for external actors.

5. CONCLUSION

This study examines the economic and political factors that influence the quality of budgetary institutions in Africa. Drawing on various econometric analyses, we found the most important factors to be control of corruption, level of voice and accountability, political competitiveness, and exposure to foreign aid. This means political factors play a more crucial role in the viability of budgetary institutions than economic factors. This is unsurprising, given the role budgetary institutions play in policy formulation and as an instrument for political settlement. The study acknowledges that due to the complex nature of governance, particularly in Africa, high levels of political will are required for budgetary quality to be improved.

These findings suggest both opportunities and challenges for improving the quality of budgetary institutions in Africa. Optimistically, political variables are largely within the purview and control of state actors and citizens, which means

achieving quality budgetary institutions relies on the political will of these actors. As Adeniran (2017) observes, effective fiscal reform depends on the preferences of the budget actors. This means achieving high-quality budgetary institutions will require alignment of interests and preferences across multiple actors. However, ensuring alignment of these disparate interests could be difficult, especially if the status quo represents a socio-political equilibrium that serves other social objectives. For example, extant institutions might be linked to a social coherence that changes might upend. Building viable and quality budgetary institutions therefore resonate with the larger discourse on the nexus between political will and governance and public sector reform in Africa.

REFERENCES

Acemoglu, D., Johnson, S., & Robinson J.A. (2003). *An African success: Botswana*. *In* In Search of Prosperity, ed. D. Rodrik, pp. 80–119, Princeton: Princeton University.

Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. Crown Books.

Adeniran, A. P. (2017). *The common problem in fiscal policy making: an exploratory experimental analysis*. Doctoral dissertation, University of the Witwatersrand. Johannesburg.

Akunyili, D., Katz, M. & Duncan, A. (2013). *Fifteen years of UK partnership with Nigeria on debt management: lessons for DFID's wider approach to building capability*. The Policy Practice and The Center for the Study of Economies of Africa. Retrieved from: https://thepolicypractice.com/wp-content/uploads/2014/03/Nigeria-Debt-Management-Report.pdf

Alesina, A. & Ferrara, E.L. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature, 43*(3), 762–800.

Alesina, A., Hausmann, R., Hommes, R. & Stein, E. (1999). Budget institutions and fiscal performance in Latin America. *Journal of Development Economics*, 59(2), 253–273.

Alesina, A. & Perotti R. (1995). The political economy of budget deficits. *IMF Staff Papers, 42*(1), 1–31.

Alesina, A. & G. Tabellini (2008). Why is fiscal policy often procyclical? *Journal of the European Economic Association*, *6*(5), 1006–1036.

Alt, J. E. & D. D. Lassen (2006). Fiscal transparency, political parties, and debt in OECD countries. *European Economic Review, 50*, 1403–1439.

Asongu, S. A. (2015). Institutional benchmarking of foreign aid effectiveness in Africa. *International Journal of Social Economics, 42*(6), 543–565. doi:10.1108/ijse-12-2013-0286.

Asongu, S.A., Nnanna, J. & Acha-Anyi, P.N. (2020). Finance, inequality and inclusive education in sub-Saharan Africa. *Economic Analysis and Policy, 67*(C), 162–177.

Asongu, S.A. & Odhiambo, N.M. (2020). Information asymmetry and insurance in Africa. *Journal of African Business, 23*(3), 394–410.

Ayee, J. R. (2008). *Reforming the African Public Sector. Retrospect and Prospects.* African Books Collective.

Bertocchi, G. & Guerzoni, A. (2010). Growth, history, or institutions: What explains state fragility in sub-Saharan Africa? *Journal of Peace Research, 49*(6), 769–783.

Brautigam, D. & Knack, S. (2004). Foreign aid, institutions, and governance in sub-Saharan Africa. *Economic Development and Cultural Change, 52*(2), 255–285.

Cabrales, A. & Hauk, E. (2011). The quality of political institutions and the curse of natural resources. *The Economic Journal, 121*(551), 58–88.

Dabla-Norris, E., Allen, A., Zanna, L., Prakash, T., Kvintradze, E. Lledo, V., Yackovlev, I. & Gollwitzer, S. (2010). *Budget Institutions and Fiscal Performance in Low-Income Countries*. IMF Working Paper WP/10/80.

Do, T. & Levchenko, A., (2009). Trade, inequality, and the political economy of institutions. *Journal of Economic Theory, 144*, 1489–1520.

Easterly, W. (2003). Can foreign aid buy growth. *Journal of Economic Perspectives*, 17(3), 23–48.

Ekeocha, D.O. & Iheonu, C.O. (2021). Household-level poverty, consumption poverty thresholds, income inequality and quality of lives in sub-Saharan Africa. *African Development Review*, 1–15.

Gollwitzer, S. (2010). Budget institutions and fiscal performance in Africa. *Journal of African Economies*, *20*(1), 111–152.

Gradstein, M. (2008). Institutional traps and economic growth. *International Economic Review*, *49*(3), 1043–1066

Hayashi, F. (2000). *Econometrics*. Princeton University Press: Princeton.

Henisz, W. J. (2004) Political institutions and volatility. *Economics and Politics, 16*, 1–27.

Iheonu, C.O. & Ichoku, H.E. (2021). Poverty and terrorism in Africa: understanding the nexus based on existing levels of terrorism. *Poverty & Public Policy, 13*(3), 254–272.

IMF (2009). *Debt limits in Fund-Supported Programs: Proposed New Guidelines*. IMF, Washington, D.C. Available at: https://www.imf.org/external/np/pp/eng/2009/080509.pdf

Islam, R., and Montenegro, C. (2002). *What determines the Quality of Institutions?* Policy Research Working Paper, World Bank.

Islam, M. (2003). Political regime and the effect of foreign aid on economic growth. *Journal of Developing Areas, 37*, 35–53.

Mahmalat, M. and Curran, D. (2018). Do crises induce reform? A critical review of conception, methodology and empirical evidence of the 'crisis hypothesis'. *Journal of Economic Surveys, 32*(3), pp.613–648.

Mulas-Granados, C. (2003). The political and economic determinants of budgetary consolidation in Europe. *European Political Economy Review, 1*(1), 015–039.

Nathaniel, S.P. and Iheonu, C.O. (2019). Carbon dioxide abatement in Africa: the role of renewable and non-renewable energy consumption. *Science of the Total Environment, 679*, 337–345.

Okwoche, P.U. and Iheonu, C.O. (2021). Determinants of fiscal effort in sub-Saharan African countries: Does conflict matter? *Economics and Business Review, 7*(2), 26–49.

Persson, T. and Tabellini G. (2000) *Political Economics, Explaining Economic Policy*. Cambridge, MA: MIT Press.

Piatti-Fünfkirchen, M. and Smets, L. (2019). *Public Financial Management, Health Financing and Under-Five Mortality. A Comparative Empirical Analysis*. IDB Working Paper Series No IDB-WP-976.

Rodrik, D., (2000). Institutions for high-quality growth: what they are and how to acquire them. *Studies in Comparative International Development, 35*(3): 3–31.

Sachs, J. D., & Warner, A. M. (2001). The curse of natural resources. *European Economic Review*, *45*(4–6), 827–838.

Schick, A. (1998). *A Contemporary Approach to Public Expenditure Management*. World Bank Institute.

Velasco, A. (2000). Debts and deficits with fragmented fiscal policymaking. *Journal of Public Economics, 76*(1), 105–125.

Von Hagen, J. (2002). Fiscal rules, fiscal institutions, and fiscal performance. *Economic and Social Review, 33*(3), 263–284.

Weingast, B., Shepsle, K., & Johansen, C. (1981). The political economy of benefits and costs: A neoclassical approach to distributive politics. *Journal of Political Economy, 89*(4), 642–664.

World Bank. (2013). *Beyond the Annual Budget: Global Experience with Medium Term Expenditure Frameworks*. Washington, DC: World Bank (doi:10.1596/978-0-8213-9625-4).

World Bank (2019). *CPIA quality of budgetary and financial management rating (1=low to 6=high)* [Data file]. Retrieved from: https://data.worldbank.org/indicator/IQ.CPA.FINQ.XQ

Received: March 02, 2021
Accepted: January 15, 2022

APPENDIX

Table 1A: Variable description and data sources

Variable	Description	Source
External debt stock	Total external debt to GDP	World Bank (Fiscal Space Database)
GDP	Real GDP	World Bank (World Development Indicator)
GDP growth	Real GDP growth	World Bank (World Development Indicator)
Trade	Imports plus exports as % of GDP	World Bank (World Development Indicator)
Aid	Development assistance and official aid received as % of GDP	World Bank (World Development Indicator)
Natural resources rents	Total natural resources rents (% of GDP)	World Bank (World Development Indicator)
Fractionalisation	Average of ethnic, linguistic, and religious fractionalisation, taken from Alesina and Ferrara (2005)	Alesina and Ferrara (2005)
System of governance	0=Presidential; 1=Parliamentary	World Bank (Database on Political Institution)
Political competition	Share of the vote of the ruling party (the variable takes 1 for non-democratic states)	World Bank (Database on Political Institution)
Voice and accountability	Country score in World Bank's Voice & Accountability Index	World Bank (World Governance Index)
Control of corruption	Country score in World Bank's Corruption Index	World Bank (World Governance Index)

Source: Authors compilation.

Table 2A: List of Countries.

Angola	Chad	Ghana	Mozambique	Sierra Leone
Benin	Congo, Democratic Republic	Kenya	Niger	Sudan
Burkina Faso	Congo	Lesotho	Nigeria	Tanzania
Cameroon	Cote d'Ivoire	Madagascar	Rwanda	Togo
Cape Verde	Ethiopia	Malawi	Sao Tome and Principe	Uganda
Central African Republic	Gambia	Mali	Senegal	Zambia
Zimbabwe				

Source: Authors compilation.

ECONOMIC ANNALS, Volume LXVII, No. 232 / January – March 2022
UDC: 3.33 ISSN: 0013-3264

https://doi.org/10.2298/EKA2232153D

*Ayşegül Durucan**

TESTING THE VALIDITY OF THE BARS CURVE FOR TURKEY

ABSTRACT: *This study has two main aims: to test the validity of the BARS curve in Turkey for the period 1974–2016 and to estimate the optimal government size for that period and compare it with the current situation. It uses the Autoregressive Distributed Lag bounds test and quadratic equation methods. The empirical findings of the study confirm the validity of the BARS curve by providing strong evidence for the existence of an inverted U-shaped long-run relationship between government size and economic growth. Unlike many previous studies that use a single proxy measure for government size, this study uses all available fundamental indicators and their sub-components. The empirical results show that for the period studied all proxy measures of government size exceed the optimal except for total central government budget expenditure and defence expenditure. Therefore, decreasing the size of the government, other than for these two indicators, will increase economic growth in the long run.*

KEY WORDS: *government size, economic growth, ARDL bounds test, Turkey*

JEL CLASSIFICATION: H11, H20, H50, O40, C19

* Assist. Prof., Kırıkkale University, Department of Public Finance, Kırıkkale, Turkey
E-mail: ayseguldurucan@gmail.com, ORCID: 0000-0001-8424-4018

** This paper is extracted from the author's PhD thesis submitted to the Ankara Yıldırım Beyazıt University, Institute of Social Sciences, Department of Public Finance.
Acknowledgements: The author is grateful to her thesis advisor Prof. Hüseyin Şen for his valuable comments and suggestions.

1. INTRODUCTION

Although the relationship between government size and economic growth has long been discussed in economic and public finance literature, there is still no consensus on whether there is a significant relationship between these variables. In addition, those that argue that there is a relationship have different approaches to determining whether this relationship is linear, non-linear, positive, or negative. The BARS curve (after the initials of its originators Barro, Armey, Rahn, and Scully) is one of these approaches.

The BARS curve can be briefly defined as showing that increases in government expenditure will increase economic growth up to a certain point but will tend to decrease growth after that point. This issue has attracted great interest in both academic and policymaking circles due to its importance in terms of public policy, resource allocation, and deciding the composition of expenditures. The many studies on this subject have yielded different results, mainly due to a lack of consensus regarding the variable(s) used as proxies for government size.

Choosing an appropriate proxy measure for government size is critically important for accurately estimating the growth-maximizing government size. To obtain reliable results the main criterion for selecting a proxy measure or measures would be a rational assessment of the specific characteristics of a country. If the proxy for government size is not chosen correctly the results will be misleading, and the policy recommendations will not result in economic growth. Therefore, both the estimation of the growth-maximizing government size and the proxy measure of the government size chosen for this estimation have been the subject of intense discussion in the literature. In contrast to many previous studies that consider a single proxy measure, this study uses all the main indicators of government size whose data can be accessed, as well as their sub-components.

Accurate determination of the relationship between government size and economic growth is a very important issue in terms of rational task sharing between public and private sector. As there is a shortage of resources in developing countries it is especially important to use existing resources effectively. It is more difficult for developing countries to recover from incorrectly allotted expenditures than it is for developed countries. Estimating the optimal

size of government to maximize growth is crucial for both task-sharing between public and private sectors and for effective public policymaking in relation to public expenditure and taxes. This paper contributes to the discussion by estimating the optimal government size for a developing country, Turkey, using 11 alternative proxy measures and comparing the obtained values with current values.

The remainder of the paper proceeds as follows. Section 2 provides the theoretical background and section 3 reviews the relevant literature. Section 4 describes the data set, methodology, and models. Section 5 presents and explains the results obtained, while Section 6 presents a robustness check of the results. Section 7 concludes.

2. THEORETICAL DISCUSSIONS ON THE BARS CURVE

The linkage between government size and economic growth is controversial. As the size of a government increases, more resources are allocated to the public sector. The government becomes the most powerful economic agent in an economy and very important for growth (Sala-i-Martin, 2002: 70). Three main mechanisms show that the positive effects of increased government size on economic growth gradually diminish and eventually become negative (Gwartney et al., 1998: 3).

Raising taxes and/or borrowing to finance increased government expenditure hurts the economy. As taxes increase the private sector's motivation to invest and increase productivity decreases. Thus, as resources are shifted from the private to the public sector, even if the efficiency of government expenditure is not diminished the disincentive effects of taxation and the crowding-out effect of domestic borrowing will harm economic growth.

As the public sector becomes larger than the private sector the diminishing returns law becomes more important. When the public sector is working properly it is expected to provide an infrastructure for the productive functioning of the private sector that includes protecting property rights, ensuring an impartial legal system, developing a stable monetary framework, and ensuring national defence, thus increasing economic growth. However, if the public sector passes its optimal growth point, government expenditure will be directed towards increasingly less

productive activities, resulting in diminishing returns and slowing economic growth.

The public sector is much less dynamic than the private sector and changes in the public sector take place much more slowly. Compared to the private sector, the time required to identify problems and to adapt to changing conditions and new knowledge and technologies is longer. This means missing opportunities for better and more efficient production. Therefore, an increase in the size of the government slows down economic growth after a threshold point.

Barro's (1990) influential paper explains the impact of government size on growth through the public policy model, a specification of endogenous growth models. According to Barro (1990), who accepts public policies as an input of production, government expenditure financed by taxes has a positive impact on economic growth up to a certain point of efficiency. However, after this effect exceeds the optimal level it starts to show negative effects on growth and leads to an inverse U-shaped association between government size and economic growth. The curve corresponding to this process is called the Barro curve and determines optimal government size. Armey (1995) also studied the relationship between government size and economic growth. The Armey curve is an inverse U-shaped non-linear curve that reveals the optimal government size by developing a quadratic function.

Several studies, including Barro (1990), King and Rebelo (1990), Jones et al. (1993), and Devarajan et al. (1996), developed endogenous growth models that incorporate fiscal policy instruments (i.e., taxes and government expenditure), assuming that such public sector policies are among of the drivers of economic growth. Barro (1990) was the most influential study on the relationship between government size and growth within the endogenous growth model framework and inspired many researchers. Armey (1995) reiterates the existence of an inverse U-shape, corresponding to a non-linear linkage between government size and growth. This non-linear linkage or 'Armey curve' shows that increased government expenditure will increase economic growth to a certain level, beyond which it will tend to lower growth. Growth-enhancing/productive government expenditure increases to a level which Armey determines as the optimal government size for maximizing growth. Increased government expenditure after

this level will lead to the law of diminishing returns and have negative consequences for economic growth.

Following Armey (1995), empirical studies by Rahn and Fox (1996) and Scully (1998, 2003) further advanced the literature on the relationship between government size and economic growth, and the curve became known the BARS curve, taking the initial letters of the surnames of Barro, Armey, Rahn, and Scully. The horizontal axis of this curve represents the government-expenditure-to-GDP ratio or government size indicator, while the vertical axis represents the real GDP growth rate. The BARS curve estimates the optimal government size to maximize economic growth and thus determines the share of the public and private sectors in a national economy and the rational policies to reach this share.

3. EMPIRICAL LITERATURE ON THE BARS CURVE

The determination of government size is a global issue. The relationship between government size and economic growth is important for determining countries' macroeconomic policies. Although a vast number of empirical studies exist, there is no consensus on the effect of government size on growth.

The empirical literature includes single-country studies that explore the optimal government size for developing countries. Rezk (2005) explores the growth-maximizing level of government size for Argentina for the period 1993–2003 and finds the optimal level of government expenditure to be as high as 30% of GDP. Herath (2009) analyses the optimal government size for Sri Lanka for the period 1959–2003 and estimates it as 27% of GDP. Abounoori and Nademi (2010) investigate the optimal government size for Iran during the period 1956–2006 and find it to be 34.7%. Alimi (2014) researches the optimal government size in Nigeria for the period 1970–2012 and concludes that it should be 19.81%. Using quarterly data for Brazil from 2000:1 to 2013:3, a more recent study by de Mendonça and Cacicedo (2015) applies the OLS and generalized method of moments (GMM) tests to Brazil for the period 2000–2013 and concludes that an increase in government size contributes positively to growth and that the optimal size for the Brazilian government is roughly 22% of GDP. A study by Şen and Kaya (2019) on Turkey estimates the optimal government size using 14 different proxy measures for the period 2006:1–2016:2. They show that the optimal

government size varies with the proxy measure used. El Husseiny (2019) examines the optimal government size for Egypt over the period 1981–2015 and finds that the value should be between 30.5% and 31.2% of GDP.

Other empirical single-country studies examine the optimal government size for developed countries. Grossman (1987) explores the optimal government size for the USA during the period 1929–1982 and concludes that it is around 19% of GDP. Peden (1991) analyses the growth-maximizing level of government size for the same country for the 1929–1986 period and finds it to be in the range of 17%–20% of GNP. Scully (1994) asserts that it should be between 21.5% and 22.9% of GNP for the USA. Likewise, Gwartney et al. (1998) explore the growth-maximizing level of government size for the USA during the 1960–1990 period and conclude that it is lower than 20% of GDP.

Chao and Grubel (1998) examine the growth-maximizing government size for Canada for the 1929–1996 period and find it to equal about 34% of national income. Vedder and Gallaway (1998) analyse the growth-maximizing government size for the USA during the period 1947–1997 and find it to be 28.87% of GDP. A study by Mittnik and Neumann (2003) examines the optimal government size for West Germany for the period 1968–1994 and concludes that it should be 20% of GDP. Mavrov (2007) examines the optimal government size for Bulgaria for the 1990–2004 period and concludes that it should be 21.4%. Scully (2008) explores the growth-maximizing level of government size for the USA during the 1960–1990 period and concludes that the ratio should be 19.3% of GDP. For France, Facchini and Melki (2013) analyse the growth-maximizing government size for the period 1896–2008 and estimate it to be around 30% of GDP. Magazzino (2014) and Forte and Magazzino (2016) analyse the relationship between government size and economic growth in Italy using a very long-term data set from 1861 to 2008 and find a non-linear relationship between the two. Di Liddo et al. (2018) investigate the relationship between government size, decentralization, and economic growth using a panel dataset for 20 Italian regions over the period 1996–2009 and find that the optimal government size is around 52%. A more recent study by Forte and Magazzino (2018) examines the optimal government size for Italy for the period 1961–2011 and report this value to be 20.6% of GDP. Makin et al. (2019) use data from 1970:1 to 2017:3 for Australia to investigate optimal government size and conclude that it should be 31%.

Other studies estimate the optimal government size in multiple countries. Karras (1996) examines the optimal government size for 118 countries over the period 1960–1985 and finds it be 23% of GDP. Karras (1997) studies the growth-maximizing level of government size for 20 European countries over the 1950–1990 period and finds it to be roughly 16% of GDP (+/–3%). Afonso et al. (2003) investigate the optimal government size for 23 OECD countries for the period 1990–2000 and conclude it should be 35%. Pevcin (2004) analyse the growth-maximizing level of government size for 12 European countries during the period 1950–1996 and finds that it to be between approximately 36% and 42% of GDP. Günalp and Dinçer (2005) research the growth-maximizing level of government size for 20 transition countries during the period 1990–2000 and estimate it to be 17.3% (+/–3%). Chobanov and Mladenova (2009) explore the optimal government size for 28 EU countries during the period 1970–2009 and conclude that it should be 25%. Forte and Magazzino (2011) explore the optimal government size for 27 EU countries during the 1970–2009 period and find that it should be between 35.39% and 43.50%. Christie (2014) examines the optimal government size for 136 countries for the 1971–2005 period and concludes that it should be 35% of GDP. Hok et al. (2014) explored the optimal government size for 8 Asian countries for the period 1995–2011 and conclude that it should be 28.50% of GDP. A more recent study by Amgain (2017) scrutinizes the optimal government size for 32 Asian countries over the period 1991–2012 and finds this value to be 18%.

Studies by Asimakopoulos and Karavias (2016) and Kim et al. (2018) show that the effect of government size on economic growth varies according to the group of countries examined. Asimakopoulos and Karavias (2016) implement three different econometrical methods – non-linear panel, GMM, and dynamic panel threshold estimation – for 129 developing and developed countries for the period 1980–2009. They observe that for developing countries the negative impact on economic growth is more pronounced in governments above optimal size than in those below the optimal size, and that the optimal government size is different for developing and developed countries: 19.12% of GDP for developing countries, 17.96% for developed countries, and 18.04% for the total of studied countries. Kim et al. (2018) examine the positive and negative effects of government size on economic growth for different threshold levels in multiple developing and developed countries. Using panel smooth transition regression (PSTR) to study

government size in 47 developing and developed countries during 1984–2012, they conclude that good governance supports government size, increases productivity, and produces economic growth; while larger government supports governance, increases productivity, and hence ensures economic growth. However, government size becomes detrimental to growth above a threshold level of government size.

4. DATA, MODEL, AND METHODOLOGY

4.1. Data

This paper uses annual time series data for Turkey spanning 1974 to 2016, which corresponds to 43 observations and is the largest timespan available. The expenditure variables used are based on the functional classification of government expenditure, thus allowing us to analyse and assess government expenditures according to their function. All variables are expressed as a percentage of GDP. Central government data are used throughout the study. The definitions of the variables and the data sources are presented in Table 1.

Descriptive statistics of the main variables are presented in Table A1 in the Appendix. Visual representations of the series are presented in Figure 1, and the justification of the variables is reported in Section 4.2.

Table 1: Definition of variables and data source

Variable		Abbrev.	Definition	Data Source
Dependent Variable		**RGdpG**	Annual Percentage Change in real GDP over the previous year *	Presidency of Strategy and Budget Office
Independent Variables	Alternative Interest Variables (Alternative Proxy Variables for Government Size)	**ECenGov**	Budget Expenditures**	General Directorate of Budget and Fiscal Control (GDBF)
		EE	Education Expenditures**	GDBF
		HE	Health Expenditures**	GDBF
		EHE	Total Health and Education Expenditures**	GDBF
		DE	Defence Expenditures**	GDBF
		CE	Final Consumption Expenditures	World Bank (WB)
		IE	Investment Expenditures	GDBF
		CIE	Total and Real Expenditures**	WB, GDBF
		RCenGov	Budget Revenues	GDBF
		RDirT	Direct Tax Revenues	GDBF
		RIndT	Indirect Tax Revenues	GDBF
	Control Variables	**DCPS**	Domestic Credit to Private Sector as a percentage of GDP is a proxy for financial sector development	WB
		INF	Inflation Rate Change in consumer price index (CPI, %)	WB
		FDI	Foreign Direct Investment (Inflows) GDP (%)	WB
		EMP	Employment Participation Rate (%) (15+)	Bulutay (1995), OECD
		CD	Crisis dummy for economic crises in the years 1980, 1994, 1999, 2001, 2009.	

* http://www.sbb.gov.tr/ekonomik-veriler/#1540461995857-3570233a-09e6

** Data for 1974–1983 is calculated by the author using Realizations of Government Expenditures and Revenues (1924–1995), Revised 2nd Edition. Data for 1983–2016 is taken from annual budget justifications.

Figure 1: Visual representation of the series from 1974 to 2016

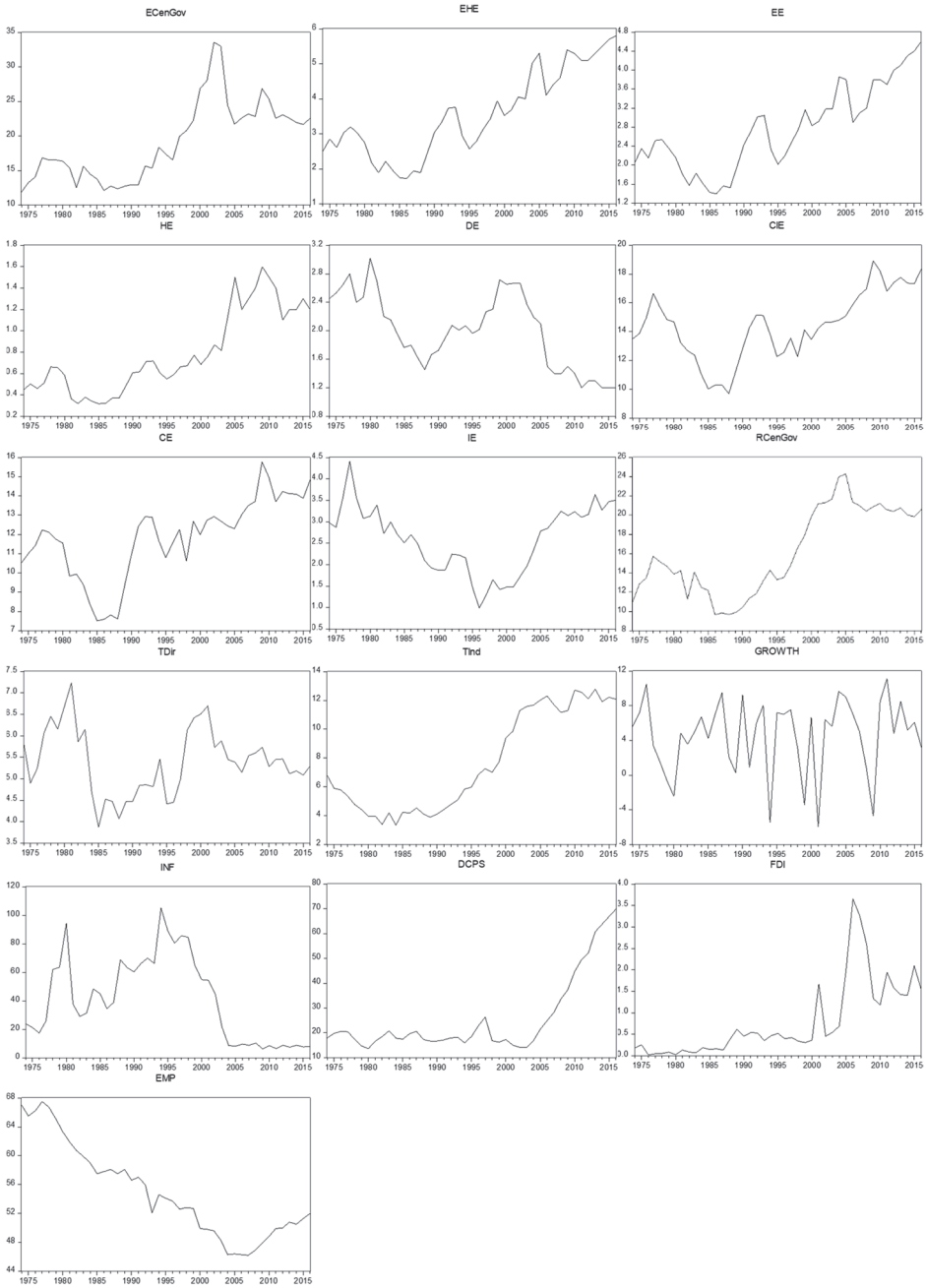


Figure 1 shows that RGdpG is affected negatively and faced sharp declines particularly in the pre-1994 crisis period, 1999 bottleneck, 2001 crisis, and 2008/09 global crisis. ECenGov has an increasing trend about until 2002, and is especially negatively affected by the 2001 crisis. It also has an increasing trend from 2005 to 2009 and is negatively influenced by the 2008/09 global crisis. The CIE decreased from the early 1980s to the early 1990s due to liberal policies, after which, except for the crisis years, it shows an increasing trend in general terms. EHE has a mostly increasing trend, but is especially negatively affected by the 1994, 2001, and 2008 crises. DE decreased from the 1980s to the beginning of the 1990s with participation in the NATO alliance and the commencement of military aid. From the 1990s to the 2000s it has an increasing trend due to the increase in terror incidents, but after 2002 it tends to decrease again. RCenGov starts to increase from 1985 due to the implementation of Value Added Tax (VAT) and peaks with the contribution of privatisation revenues in 2005, but after 2005 it decreases until 2008. RDirT, the share of direct taxes, decreased with the effect of liberalisation after 1980. RIndT started to increase from the beginning of 1990 due to the introduction of VAT, like RCenGov.

4.2. Model

The control variables are (1) inflation rate (INF), (2) domestic credit to private sector (DCPS), (3) foreign direct investment (FDI), (4) employment participation rate (EMP), and (5) a dummy variable for the economic and political crisis. To eliminate possible seasonality problems in prices, INF refers to annual changes in the consumer price index over the previous year. Empirical studies in the literature have produced mixed results regarding the linkage between inflation and growth. Barro (1996), Andres and Hernando (1999), and Gillman and Harris (2010) find a negative effect of inflation on growth, while Kormendi and Meguire (1985) and Mallik and Chowdhury (2001) report the opposite result in the long run. Considering these studies, we posit that the link between inflation and growth may be positive or negative.

The second variable is domestic credit to the private sector (DCPS), which is widely accepted as a proxy for financial sector development. DCPS plays a very important role in increasing investment and employment in an economy, not only providing efficiency and productivity but also enhancing economic growth (Begum and Aziz, 2019). Bencivenga et al. (1996) conclude that the development

of the financial sector through increased liquidity positively affects economic growth. On the other hand, it has been argued that financial sector development is negatively related to economic growth. For example, Olowofeso et al. (2015: p.82) point out that "the efficient provisioning of credit has a positive and significant effect on output and employment opportunities, while a low level of financial development and its attendant inefficient private sector credit system distorts economic growth". In this paper the expected sign of financial sector development is negative, since Turkey is a developing country with an insufficiently developed and illiquid financial market and there are limited opportunities for it to significantly impact economic growth.

The third control variable is foreign direct investment (FDI). The literature on the FDI–growth nexus has yielded mixed results. Borensztein et al. (1998) find that FDI is a significant mechanism for the transfer of technology and its contribution to economic growth is more positive than domestic investment. On the other hand, there are some situations in which FDI may adversely affect the host country's economy. The high competitiveness of the investing country can adversely affect companies in the host country and reduce their market share (Kahveci and Terzi, 2017). However, when assessing the impact of FDI on growth the type of investment should be taken into account. FDI can be realised as new investment, acquisitions, or mergers. If FDI is realised as an acquisition or merger it is expected that the investment will create a new production area and will have a positive effect on growth. Hence, there is no agreed consensus on the sign of this variable: FDI can enter the model with a positive or negative sign.

Our fourth control variable is the employment participation rate (EMP). Only a few studies in the literature examine the association between employment and economic growth, and they show mixed results. Saget (2000) finds mixed results for different countries: the relationship between employment and economic growth has a positive effect on economic growth in Hungary, Poland, the Czech Republic, Slovakia, and Russia, a negative impact in Romania, and no effect in Bulgaria and Ukraine. Yam et al. (2002) find the relationship to be positive, whereas Abdullah et al. (2011) find no significant link between the variables. Thus, there is no consensus about the sign of the employment participation rate variable.

Our fifth control variable is the crisis dummy. This dummy variable posits the years with negative economic growth as crisis years. The crisis years for the Turkish economy are 1980, 1994, 1999, 2001, and 2009. This dummy variable removes the effect of the crisis from the effect of government size on growth. Thus, the expected sign of the crisis dummy in this paper is negative.

Our sixth and main variable is government size. In the literature there are different proxy measures for government size. However, the ratio of government expenditure to GDP is the most commonly used proxy measure.

The first eight models (Model 1 through Model 8 that are presented in below) are based on the government expenditures while further three models (Model 9 through Model 11) are based on government revenues.

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, EcenGov, EcenGov^2) \quad (1)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, EE, EE^2) \quad (2)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, HE, HE^2) \quad (3)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, EHE, EHE^2) \quad (4)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, DE, DE^2) \quad (5)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, CE, CE^2) \quad (6)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, IE, IE^2) \quad (7)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, CIE, CIE^2) \quad (8)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, RCenGov, RCenGov^2) \quad (9)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, RdirT, RdirT^2) \quad (10)$$

$$RGdpG = f(INF, DCPS, FDI, EMP, CD, RindT, RindT^2) \quad (11)$$

4.3. Methodology

To research the presence of a possible long-run linkage between government size and growth, first, the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller 1979) and Phillips-Perron (PP) test (Phillips and Perron 1988) are used to determine the appropriate estimation technique. These unit root tests are used to scrutinize the stationarity properties of the level and first difference of the variables. Second, the cointegration relationship between the variables is estimated using the ARDL bounds test (Pesaran et al., 2001).

5. EMPIRICAL FINDINGS

5.1. Unit root test results

The ADF and the PP unit root tests were applied to empirically examine the stationarity of all the variables. The ADF and the PP tests of the null hypothesis of non-stationarity are tested against the alternative of stationarity. The results of the tests are represented in Tables 2 and 3. ECenGov, EHE, EE, HE, DE, CIE, CE, IE, RCenGov, RdirT, RindT are stationary at the first difference I(1) for both models. However, RGdpG is stationary at level I(0) for both models, FDI is stationary at level I(0) for the constant and trend model, and RdirT is stationary at level I(0) for the constant model. Thus, the results show that the series are stationary at different orders. Because the variables have different integration orders and there is no I(2) data, we can safely use the ARDL bounds test.

Table 2: Unit root test results, 1974–2016

Series	ADF unit root test						PP unit root test					
	Level	Critical Values		Level	Critical Values		Level	Critical Values		Level	Critical Values	
	Constant	5%: –2.93	1%: –3.59	Constant and Trend	5%: –3.52	1%: –4.19	Constant	5%: –2.93	1%: –3.59	Constant and Trend	5%: –3.52	1%: –4.19
RGdpG	–6.26(0)***			–6.23(0)***			–6.26***			–6.29***		
CE	–1.10(0)			–2.06(0)			–1.32			–2.30		
IE	–1.28(0)			–1.20(0)			–1.36			–1.10		
CIE	–0.83(0)			–2.03(0)			–1.15			–1.84		
HE	–1.00(0)			–2.19(0)			–0.94			–2.24		
EE	–0.34(0)			–2.81(1)			–0.20			–2.21		
EHE	–0.45(0)			–2.28(0)			–0.19			–2.17		
DE	–1.02(0)			–1.65(0)			–1.20			–1.95		
ECenGov	–1.59(0)			–2.24(1)			–1.66			–2.07		
RCenGov	–1.04(0)			–2.04(2)			–1.12			–1.59		
RdirT	–2.52(0)			–2.48(0)			–2.62*			–2.59		
RindT	0.11(0)			–2.82(0)			–0.18			–2.77		
FDI	–1.94(0)			–3.65(1)**			–1.82			–2.69		
DCPS	2.79(0)			0.74(0)			2.41			0.55		
INF	–1.61(0)			–2.18(0)			–1.56			–2.09		
EMP	–2.05(0)			–0.21(0)			–2.03			–0.06		

Note: The numbers in parentheses indicate the chosen lag order of the ADF models. Lags are selected based on AIC. Asterisks (*), (**), (***) denote statistical significance at 10%, 5%. and 1%, respectively. E-Views 10 was used for computations.

Table 3: Unit root test results, 1974–2016

ADF unit root test							PP unit root test					
Series	**First Difference Constant**	**Critical Values**		**First Difference Constant and Trend**	**Critical Values**		**First Difference Constant**	**Critical Values**		**First Difference Constant and Trend**	**Critical Values**	
		5%:–2.93	**1%: –3.60**		**5%: –3.52**	**1%: –4.20**		**5%: –2.93**	**1%: –3.60**		**5%: –3.52**	**1%: –4.19**
D(RGdpG)	–7.11(1)***			–7.04(1)***			–19.81***			–19.64***		
D(CE)	–5.98(0)***			–5.95(0)***			–5.99***			–5.96***		
D(IE)	–5.50(1)***			–5.89(1)***			–5.93***			–6.05***		
D(CIE)	–5.22(0)***			–5.26(0)***			–5.20***			–5.25***		
D(HE)	–6.22(0)***			–6.15(0)***			–6.23***			–6.15***		
D(EE)	–5.59(0)***			–4.82(1)***			–5.57***			–6.13***		
D(EHE)	–5.81(0)***			–5.11(1)***			–5.88***			–7.34***		
D(DE)	–2.81(2)*			–2.74(2)			–5.69***			–5.62***		
D(EcenGov)	–5.58(0)***			–5.52(0)***			–5.58***			–5.52***		
D(RcenGov)	–6.35(0)***			–6.26(0)***			–6.35***			–6.26***		
D(RdirT)	–7.00(0)***			–6.92(0)***			–6.99***			–6.90***		
D(RindT)	–3.28(1)**			–3.27(1)***			–6.18***			–6.29***		
D(FDI)	–5.64(2)***			–5.55(2)***			–10.60***			–10.38***		
D(DCPS)	–4.05(0)***			–5.09(0)***			–4.00***			–5.05***		
D(INF)	–6.83(0)***			–6.91(0)***			–6.89***			–7.17***		
D(EMP)	–6.08(0)***			–6.59(0)***			–6.14***			–6.59***		

Note: The numbers in parentheses indicate the chosen lag order of the ADF models. Lags are selected based on AIC. Asterisks (*), (**), (***) denote statistical significance at 10%, 5%, and 1%, respectively. E–Views 10 was used for computations.

5.2. ARDL bounds test results

In the previous subsection, two unit root tests were implemented on both the level and the first differenced forms to capture the appropriate econometric method for estimation. The test results show a group of time series, some I(0), others I(1), but no I(2). The next step investigates the existence of a cointegration relationship among the variables using the bounds test approach (Pesaran et al. 2001). The presence of an inverse U-shaped association between the size of government and growth was estimated using the quadratic equation method, regressing the real GDP growth rate on the size of government and the government size squared. Baseline regressions were performed following Vedder and Gallaway (1998), Chao and Grubel (1998), Pevcin (2004), and El Husseiny (2019). The subsequent specification was tested for Turkey over the period 1974–2016. If the squared coefficient on the proxy measure for government size is negative and statistically significant, it means that economic growth initially increases and eventually decreases with the rise in government size. The growth-maximizing government size can be formulated as follows:

$$\text{Govsize}^{*} = \frac{-\beta_{15}}{2\beta_{16}}$$

$$\begin{aligned}\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 &+ \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} \\ &+ \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} \\ &+ \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta \boldsymbol{E_{CenGov}}_{t-i} \\ &+ \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta \boldsymbol{E_{CenGov}}^{2}{}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} \\ &+ \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} \\ &+ \beta_{15} \boldsymbol{E_{CenGov}}_{t-1} \\ &+ \beta_{16} \boldsymbol{E_{CenGov}}^{2}{}_{t-1} \\ &+ \mu_t \end{aligned} \quad (1)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta EE_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {EE^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} EE_{t-1} + \beta_{16} {EE^2}_{t-1} + \mu_t \quad (2)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta HE_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {HE^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} HE_{t-1} + \beta_{16} {HE^2}_{t-1} + \mu_t \quad (3)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta EHE_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {EHE^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} EHE_{t-1} + \beta_{16} {EHE^2}_{t-1} + \mu_t \quad (4)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta DE_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {DE^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} DE_{t-1} + \beta_{16} {DE^2}_{t-1} + \mu_t \quad (5)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta CE_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {CE^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} CE_{t-1} + \beta_{16} {CE^2}_{t-1} + \mu_t \quad (6)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta IE_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {IE^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} IE_{t-1} + \beta_{16} {IE^2}_{t-1} + \mu_t \quad (7)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta CIE_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {CIE^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} CIE_{t-1} + \beta_{16} {CIE^2}_{t-1} + \mu_t \quad (8)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta \boldsymbol{R_{CenGov}}_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta {\boldsymbol{R_{CenGov}}^2}_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} \boldsymbol{R_{CenGov}}_{t-1} + \beta_{16} {\boldsymbol{R_{CenGov}}^2}_{t-1} + \mu_t \quad (9)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta \boldsymbol{R_{DirT}}_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta \boldsymbol{R_{DirT}}^2_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} \boldsymbol{R_{DirT}}_{t-1} + \beta_{16} \boldsymbol{R_{DirT}}^2_{t-1} + \mu_t \quad (10)$$

$$\Delta \boldsymbol{G_{RGdp}}_t = \beta_0 + \sum_{i=1}^{p-1} \beta_{1i}\, \Delta \boldsymbol{G_{RGdp}}_{t-i} + \sum_{i=0}^{p_1-1} \beta_{2i}\, \Delta Inf_{t-i} + \sum_{i=0}^{p_2-1} \beta_{3i}\, \Delta DCPS_{t-i} + \sum_{i=0}^{p_3-1} \beta_{4i}\, \Delta FDI_{t-i} + \sum_{i=0}^{p_4-1} \beta_{5i}\, \Delta Emp_{t-i} + \sum_{i=0}^{p_5-1} \beta_{6i}\, \Delta CD_{t-i} + \sum_{i=0}^{p_6-1} \beta_{7i}\, \Delta \boldsymbol{R_{IndT}}_{t-i} + \sum_{i=0}^{p_7-1} \beta_{8i}\, \Delta \boldsymbol{R_{IndT}}^2_{t-i} + \beta_9 \boldsymbol{G_{RGdp}}_{t-1} + \beta_{10} Inf_{t-1} + \beta_{11} DCPS_{t-1} + \beta_{12} FDI_{t-1} + \beta_{13} Emp_{t-1} + \beta_{14} CD_{t-1} + \beta_{15} \boldsymbol{R_{IndT}}_{t-1} + \beta_{16} \boldsymbol{R_{IndT}}^2_{t-1} + \mu_t \quad (11)$$

where Δ symbolizes the first difference, β_0 defines the intercept component, βs indicates the coefficients of variables, and μ is the error term or white noise residuals. The ARDL model estimates Equations (1) through (11) to get the optimal lag order for each variable. The subsequent hypotheses should be tested to decide determine the presence of cointegration among variables. The null hypothesis of no cointegration is as follows (Pesaran et al., 2001):

$$H_0: \begin{bmatrix} \beta_9 \\ \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \\ \beta_{15} \\ \beta_{16} \end{bmatrix} = 0_{8\times1}$$ is tested against the alternative; that is: $H_1: \beta_9 \neq 0$ or $H_1: \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \\ \beta_{15} \\ \beta_{16} \end{bmatrix} \neq 0_{7\times1}$

The time series are cointegrated if the computed F-statistics are greater than the appropriate higher bounds I(1) and not cointegrated if the computed F-statistics are below the lower bounds I(0) of Pesaran et al. (2001). Equation (1) provides the short-run and long-run effects concurrently after the adjustment is completed. The long-run effects are inferred by the estimates of β_{10}, β_{11}, β_{12}, β_{13}, β_{14}, β_{15} *and* β_{16} that are normalized on the estimate of β_9. After obtaining evidence for the presence of cointegration among variables, the optimal lag orders of each variable were chosen using the suitable AIC. According to the empirical literature,[1] the maximum lag order is usually between 2 and 4 for annual data in order not to lose more degrees of freedom, which is very important for the reliability of the results. Since in this study the data set is relatively large, the maximum lag of 3 was taken to find the cointegration among variables. Selected ARDL models in Table 4 below are revealed as optimal for the series of models used in this paper.

[1] See, for example, Pesaran and Shin, 1999; Pesaran et al. 2001.

Table 4: Selection of optimal models

	Selected Model
Model 1	ARDL(3, 3, 3, 3, 2, 3, 0, 3)
Model 2	ARDL(3, 3, 3, 3, 3, 3, 3, 3)
Model 3	ARDL(3, 3, 2, 3, 3, 3, 1, 2)
Model 4	ARDL(3, 3, 3, 3, 3, 3, 3, 3)
Model 5	ARDL(1, 1, 3, 2, 0, 2, 2, 2)
Model 6	ARDL(2, 0, 2, 2, 2, 2, 0, 2)
Model 7	ARDL(1, 0, 2, 2, 2, 0, 0, 1)
Model 8	ARDL(1, 0, 0, 2, 2, 0, 0, 0)
Model 9	ARDL(3, 3, 3, 3, 3, 3, 3, 0)
Model 11	ARDL(2, 0, 2, 2, 0, 0, 2, 1)
Model 12	ARDL(1, 0, 2, 1, 3, 3, 3, 3)

Table 5 summarizes the computed F-statistics of the models. These values are compared with the critical values of Pesaran et al. (2001). The results indicate that cointegration relationships exist between independent variables for all models.

Table 5: F-Bounds test results (Null hypothesis: No level relationship)

Test Statistic	**Value**	**Signif.**	**I(0)**	**I(1)**
Model 1			Asymptotic: n=1000	
F-statistic	**8.00**	1%	2.96	4.26
Model 2			Asymptotic: n=1000	
F-statistic	**6.16**	1%	2.96	4.26
Model 3			Asymptotic: n=1000	
F-statistic	**11.91**	1%	3.31	4.63
Model 4			Asymptotic: n=1000	
F-statistic	**3.49**	5%	2.32	3.5
Model 5			Asymptotic: n=1000	
F-statistic	**14.83**	1%	2.54	3.91

Model 6			Asymptotic: n=1000	
F-statistic	**6.68**	1%	2.54	3.91
Model 7			Asymptotic: n=1000	
F-statistic	**20.03**	1%	2.96	4.26
Model 8			Asymptotic: n=1000	
F-statistic	**25.75**	1%	3.31	4.63
Model 9			Asymptotic: n=1000	
F-statistic	**6.76**	1%	2.96	4.26
Model 10			Asymptotic: n=1000	
F-statistic	**23.73**	1%	2.54	3.91
Model 11			Asymptotic: n=1000	
F-statistic	**13.2470**	1%	2.54	3.91

Since all series are cointegrated, the long-run coefficients can be estimated. The estimated long-run coefficients of the variables are presented in Tables 6 and 7. Diagnostic tests were conducted. The autocorrelation problem was found in all models except 6, 8, and 10. To solve this problem, the Heteroscedasticity and Autocorrelation Consistent (HAC) estimator was applied. Other results provide no evidence of the diagnostic problem in the long-run estimation of all models. CUSUM and CUSUM-Q test results are given in Figure A1 in the Appendix.

Table 6: Long-run estimate results for government expenditure.

Dependent Variable: Annual Real GDP Growth Rate

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INF	−0.03***	−0.05***	−0.01*	−0.05***	−0.04***	−0.04**	−0.00	−0.02
	(−3.16)	(−4.67)	(−2.6)	(−4.16)	(−4.17)	(−2.67)	(−0.21)	(−1.23)
DCPS	−0.05***	0.08	0.36*** (5.84)	0.04	0.01	0.00	−0.01	0.40***
	(−4.30)	(1.66)		(1.24)	(0.90)	(0.04)	(−0.89)	(4.23)
FDI	−4.62***	−4.76***	−4.03***	−3.91***	−2.87***	−3.02***	−3.59***	−3.96***
	(−6.08)	(−6.75)	(−5.96)	(−7.01)	(−7.54)	(−3.42)	(−3.71)	(−3.63)
EMP	−0.34***	−0.46***	−1.80***	−0.42***	−0.22***	−0.27***	−0.22	−1.97***
	(−5.47)	(−7.12)	(−6.66)	(−7.08)	(−5.00)	(−3.52)	(−1.00)	(−4.56)
CD	−11.52***	−15.280***	−7.040***	−14.11***	−7.61***	−7.91***	−6.06***	−6.33***
	(−7.72)	(−13.299)	(−4.962)	(−13.16)	(−7.29)	(−3.63)	(−4.28)	(−4.49)
ECenGov	0.66***							
	(3.13)							
ECenGov^2	−0.01**							
	(−2.81)							
EE		11.48**						
		(2.79)						
EE^2		−2.06**						
		(−2.38)						
HE			24.43*** (5.97)					
HE^2			−10.27***					
			(−4.65)					
EHE				7.45***				
				(4.03)				
EHE^2				−1.06***				
				(−3.46)				
DE					20.75***			
					(10.01)			
DE^2					−5.00***			
					(−10.79)			
CE						5.06***		
						(4.65)		
CE^2						−0.24***		
						(−4.29)		
IE							5.48**	
							(2.06)	
IE^2							−1.06*	
							(−1.72)	
CIE								5.90**
								(2.29)
CIE^2								−0.19*
								(−2.02)

Trend			−2.96***					−1.25***
			(−4.52)					(−4.32)
Constant	58.21***	50.06***	289.74***	48.87***			14.73	99.07***
	(4.76)	(3.53)	(5.23)	(4.04)			(1.38)	(3.35)
Opt. Gov. Size	23.85	2.78	1.18	3.50	2.07	10.29	2.57	15.07
R^2	0.95	0.97	0.97	0.97	0.93	0.86	0.84	0.82
Adjusted R^2	0.84	0.88	0.92	0.87	0.86	0.74	0.74	0.74
Observations	40	40	40	40	40	41	41	41

Note: Values in parentheses are t-statistics. Asterisks (*), (**), (***) denote statistical significance at 10%, 5%, and 1%, respectively.

Table 7: Long run estimate results for government revenues
Dependent Variable: Annual Real GDP Growth Rate

Variable	**(9)**	**(10)**	**(11)**
INF	−0.06***	−0.04***	−0.00
	(−3.67)	(−3.38)	(−0.42)
DCPS	−0.07***	−0.03	0.12**
	(−4.44)	(−1.61)	(2.13)
FDI	−5.16***	−3.21***	−1.13
	(−8.14)	(−5.69)	(−1.68)
EMP	−0.50***	−0.18***	−0.09
	(−6.22)	(−3.70)	(−1.08)
CD	−11.17***	−4.49***	−25.96***
	(−7.14)	(−3.15)	(−6.28)
RCenGov	1.84**		
	(3.00)		
$RCenGov^2$	−0.05**		
	(−2.81)		
RDirT		8.77***	
		(7.83)	
$RDirT^2$		−0.93***	
		(−8.61)	
RIndT			3.31*
			(2.00)
$RIndT^2$			−0.26*
			(−2.03)
Constant	56.69***		
	(5.35)		
Opt. Gov. Size	17.10	4.68	6.34
R^2	0.95	0.86	0.93
Adjusted R^2	0.85	0.78	0.84

Note: Values in parentheses are t-statistics. Asterisks (*), (**), (***) denote statistical significance at 10%, 5%, and 1%, respectively.

Table 8. Basic statistics related to government size

Alternative Proxy Measures for Government Size	Mean	Maximum	2016 Realization Rate	Optimal Rate*	Observations
ECenGov	19.11	33.54	22.55	23.85	40
CIE	14.40	18.91	18.34	15.07	41
CE	11.80	15.77	14.84	10.29	41
IE	2.59	4.40	3.50	2.57	41
EHE	3.54	5.80	5.8	3.50	40
EE	2.75	4.60	4.60	2.78	40
HE	0.78	1.60	1.20	1.18	40
DE	1.99	3.01	1.20	2.07	40
RCenGov	16.19	24.28	20.58	17.10	40
RDirT	5.39	7.22	5.29	4.68	41
RIndT	7.68	12.77	12.08	6.34	40

* It covers 1974–2016.

Table 8 presents basic statistics related to proxy measures for government size. Because the means of the proxy measures used for government size and the 2016 realization rates are significantly different, the optimal rates obtained from the econometric analysis are compared with the 2016 realizations in order to reach more realistic inferences and enable future rational policy proposals. Figure 2 presents the BARS curve for Turkey by alternative proxy measures for government size. The realized central government budget expenditure in 2016 is 22.55% of GDP and the optimum estimated for this expenditure indicator is 23.85%, so the government size indicator is below the optimal ratio.

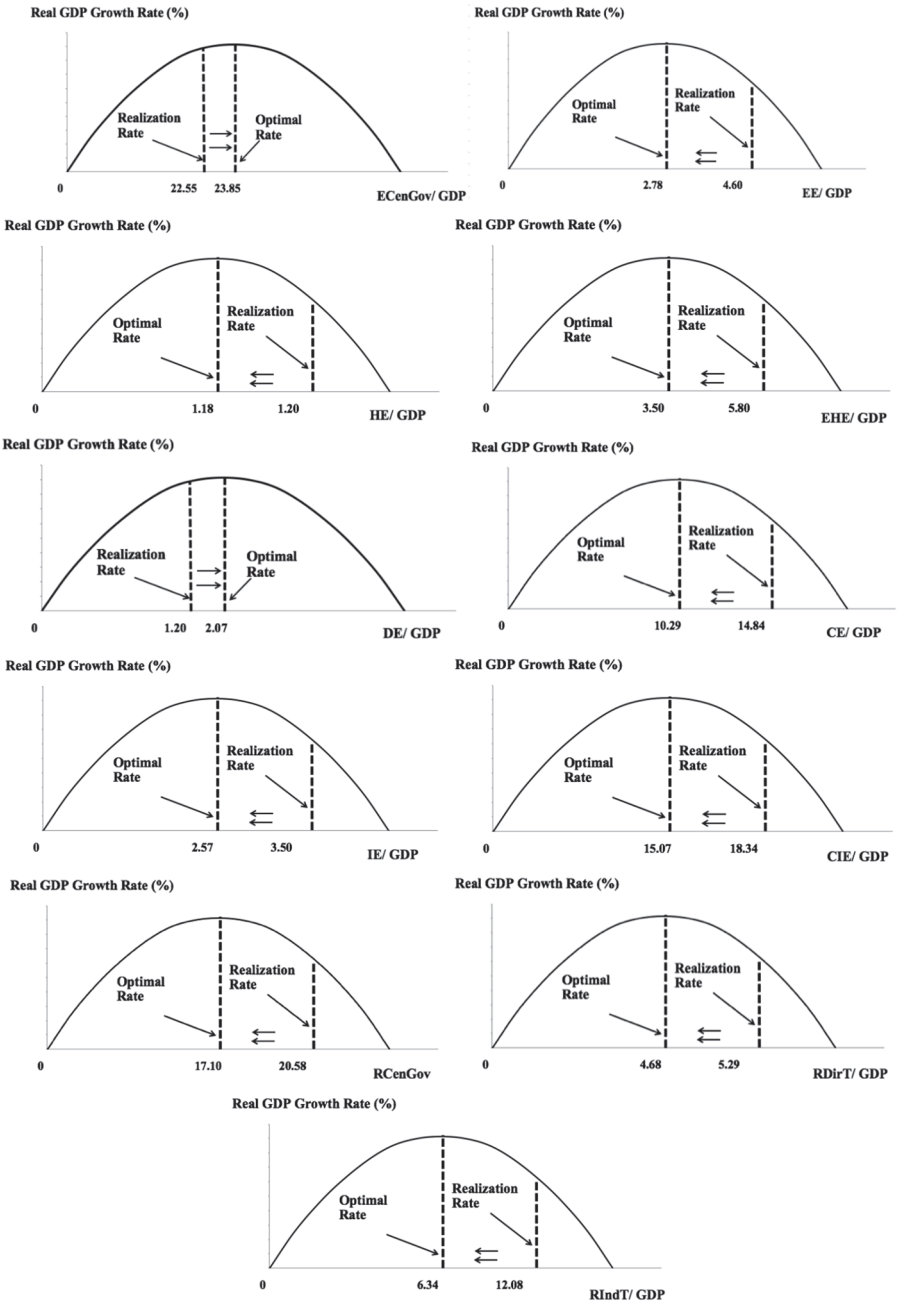
Central government real expenditure consists of the sum of central government investment and consumption expenditure. The results suggest that the central government's real expenditure should be reduced. Which component or components of this expenditure should be decreased is important, so central government investment and consumption expenditure are estimated separately. The 2016 realization rate of central government consumption expenditure for the

year 2016 is 14.84% of GDP, and the optimal size of government is estimated at 10.29%. The realization rate of central government investment expenditure in 2016 is 3.50% of GDP and the optimal rate is estimated to be 2.57%. These findings suggest that both central government consumption and investment expenditure should be reduced to avoid a negative effect on economic growth.

Additionally, the 2016 realization rates of central government education and health expenditure are 4.60% and 1.20% of GDP respectively, above the estimated optimal rates of 2.78% and 1.18%. Although the realization rate for central government education and health expenditure to GDP is well above the optimal rate, the inability to obtain the desired efficiency from the expenditure shows that it was not used efficiently and effectively. On the other hand, central government defence expenditure in 2016 is 1.20% of GDP, whereas the estimated optimal rate for this indicator is 2.07% of GDP, indicating that defence expenditure should be increased.

The examination of central government budget revenue as an indicator of government size shows that the central government budget revenue for the year 2016 is 20.58% of GDP. The estimated optimal percentage for this proxy measure is 17.10%, so the current rate is above the optimal rate. In this study the central government's direct and indirect tax revenues are estimated separately. The realization rate of the central government's direct tax revenue in 2016 is 5.29% of GDP, while the estimated optimal rate is 4.68% of GDP. The realization rate of the central government's indirect tax revenue in 2016 is 12.08% of GDP and the estimated optimal rate is 6.34% of GDP. These findings show that the central government budget revenue is well above the estimated optimal percentage of GDP for central government direct and indirect tax revenue in Turkey.

Figure 2: Bars curves for Turkey by proxy measure of government size



6. ROBUSTNESS CHECK

When performing empirical analysis, strict assumptions may be made that can affect the results. It is necessary to test whether the results of the empirical analysis are robust with different methods or variables. Therefore, a robustness test is performed to verify long-run analysis results. Central government budget revenues lagged by one year are added to the first eight and the last three models in which central government budget expenditure and its sub-components are used as proxies for government size. The variables are lagged by one year to eliminate the endogeneity problem (Wooldridge, 2012) between the central government budget expenditure and revenue. Tables 9 and 10 show that the long-run estimations confirm the inverted U-shaped relationship between government size and economic growth for Turkey.[2]

Table 9: Robustness check results for government expenditure

Dependent Variable: Annual Real GDP Growth Rate

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INF	−0.08**	−0.02	−0.00	−0.00	−0.04**	−0.04	0.01	0.00
	(−2.58)	(−0.69)	(−0.48)	(−0.27)	(−2.68)	(−1.50)	(0.41)	(0.27)
DCPS	−0.09**	0.11	0.49***	0.42***	−0.00	0.01	0.01	0.74***
	(−2.59)	(1.47)	(8.02)	(4.10)	(−0.28)	(0.44)	(0.74)	(6.92)
FDI	−3.65***	−4.52***	−3.94***	−4.68***	−3.13***	−3.34**	−3.83***	−8.82***
	(−3.88)	(−3.21)	(−8.78)	(−6.58)	(−5.65)	(−2.53)	(−4.48)	(−6.58)
EMP	−0.13**	−0.23**	−2.37***	−1.95***	−0.14***	−0.29**	−0.03	−3.68***
	(−2.86)	(−2.35)	(−8.33)	(−5.31)	(−3.02)	(−2.22)	(−0.39)	(−7.64)
CD	−	−7.89***	−5.91***	−7.21***	−8.19***	−8.07***	−7.29***	−4.96***
	17.39***	(−4.88)	(−3.73)	(−4.64)	(−9.10)	(−3.20)	(−4.32)	(−4.93)
	(−6.51)							
ECenGov	3.18***							
	(5.57)							
ECenGov^2	−0.04***							
	(−4.78)							
EE		12.93**						
		(2.52)						
EE^2		−2.33**						
		(−2.30)						

[2] Diagnostic tests for autocorrelation, heteroscedasticity, normality, and stability (Cusum and Cusum-Q) were conducted. Results from the authors upon request.

HE			35.93***					
			(4.51)					
HE²			−15.61***					
			(−3.98)					
EHE				9.89**				
				(2.80)				
EHE²				−1.16**				
				(−2.28)				
DE					16.83***			
					(5.54)			
DE²					−4.31***			
					(−8.25)			
CE						5.59**		
						(2.78)		
CE²						−0.27***		
						(−3.09)		
IE							4.84*	
							(1.82)	
IE²							−1.30***	
							(−3.19)	
CIE								6.70***
								(3.46)
CIE²								−0.22**
								(−3.14)
RCenGov(−1)	−1.27**	0.20	0.00	0.16**	0.11	−0.05	0.35***	0.74***
	(−3.26)	(1.11)	(0.05)	(2.31)	(0.88)	(−0.25)	(4.13)	(5.45)
Trend			−2.80***	−2.09***				−2.77***
			(−5.22)	(−3.97)				(−6.99)
Constant			255.05***	193.29***				232.51***
			(5.95)	(4.76)				(6.97)
Opt. Gov. Size	33.17	2.76	1.15	4.23	1.94	10.06	1.86	15.16
R^2	0.97	0.78	0.96	0.94	0.90	0.89	0.94	0.99
Adjusted R^2	0.84	0.68	0.88	0.84	0.83	0.77	0.82	0.96
Observations	39	40	40	40	41	41	40	40

Note: Values in parentheses are t-statistics. Asterisks (*), (**), (***) denote statistical significance at 10%, 5%, and 1%, respectively.

Table 10: Robustness check results for government revenue

Dependent Variable: Annual Real GDP Growth Rate

Variable	(9)	(10)	(111)
INF	–0.02	–0.01	–0.07***
	(–0.93)	(–0.85)	(–5.50)
DCPS	–0.05**	–0.02	0.07
	(–2.78)	(–0.91)	(1.04)
FDI	–2.79***	–2.10**	–4.96***
	(–3.89)	(–2.63)	(–8.12)
EMP	–0.14	0.04	–0.57***
	(–1.58)	(0.33)	(–4.92)
CD	–13.34***	–4.31***	–16.50***
	(–12.58)	(–3.98)	(–5.33)
RCenGov	1.57*		
	(1.88)		
RCenGov^2	–0.06*		
	(–2.00)		
RDirT		19.55***	
		(3.06)	
RDirT^2		–1.95***	
		(–3.36)	
RIndT			7.46***
			(6.54)
RIndT^2			–0.47***
			(–4.34)
ECenGov(–1)	0.52**	0.09	–0.21
	(2.32)	(0.89)	(–0.71)
Constant		–61.74**	61.82
		(–2.32)	(1.74)
Opt. Gov. Size	12.51	5.00	7.90
R^2	0.81	0.89	0.98
Adjusted R^2	0.68	0.80	0.92
Observations	41	41	39

Note: Values in parentheses are t-statistics. Asterisks (*), (**), (***) denote statistical significance at 10%, 5%, and 1%, respectively.

7. CONCLUDING REMARKS

Studies examining the relationship between government size and economic growth yield quite different results. The main reason for these differences is the variable or variables used as indicators of government size. Other possible causes are the specific characteristics of the country studied, the period studied, and the different econometric methods used.

In this context, the purpose of this article is twofold. First, it examines whether the inverse U-shaped relationship between government size and economic growth is valid for Turkey in the period 1974–2016. Second, it estimates the optimal government size for this period, based on various indicators, in order to compare it with the current situation. Long-term coefficients were obtained using the ARDL bounds test, then the square of the variable used for government size was added to the model. Next, the optimal government size ratios were calculated by testing whether the relationship between economic growth and government size was linear.

Our findings show that there is a long-run relationship between government size and economic growth, and this relationship confirms the validity of the BARS curve. However, this validation depends on how government size is measured. As long as the 11 variables – ECenGov, EE, HE, EHE, DE, CE, IE, CIE, RCenGov, RIndT, RDirT – are taken as proxies for government size, the results support the existence of this relationship. Moreover, our results are in line with other studies (Peden 1991; Scully 1994; Karras 1996; Vedder and Gallaway 1998; Chao and Grubel 1998; Afonso et al. 2003; Chobanov and Mladenova 2009; Forte and Magazzino 2011; Christie 2014; Asimakopoulos and Karavias 2016; Forte and Magazzino 2016; Şen and Kaya 2019) that also found a non-linear relationship between government size and economic growth.

When the empirical findings of the study are evaluated in terms of expenditure, the estimation results for the optimal government size show that expenditures other than the central government budget and defence expenditure are above the estimated optimal rates and should be reduced. A decrease in these expenditures to the calculated optimal rates will lead to increased economic growth.

The government size that maximizes economic growth in terms of taxes shows that the realization rates of both indirect and direct tax revenues are above the optimal rates. Therefore, to maximize economic growth in Turkey, both indirect and direct taxes should be reduced until the calculated optimal rates are reached.

These results show that it is possible to achieve economic growth without additional financing by changing the composition of expenditures and using existing resources more effectively and efficiently. In particular, it is vital for developing countries to implement policies to avoid wasting resources and to achieve and even maintain economic growth without the need for borrowing. Fiscal policy instruments such as public expenditure and taxes should not fall below or exceed the calculated optimal ratios.

This study has some limitations. It focuses primarily on central government expenditure and revenue and their main sub-components. However, these sub-components also have other sub-components, which could be examined in more detail in future research. The data is limited to the period 1974–2016, as no longer-term data for the variables is available. Finally, the analysis used a linear econometric method; future studies could contribute to the literature by using non-linear methods.

REFERENCES

Abdullah, N., Naim, Z. A., & Long, Y. (2011). Employment and macroeconomic variables: evidence from Malaysia, Philippines, and Singapore. *International Journal of Economics and Finance*, 3 (3), 139–148.

Abounoori, E. & Nademi, Y. (2010). Government size threshold and economic growth in Iran. *International Journal of Business and Development Studies*, 2, 95–108.

Afonso, A., Schuknecht, L., & Tanzi, V. (2003). Public sector efficiency: An international comparison. *Working Paper Series 242*. European Central Bank, Frankfurt.

Alimi, R.S. (2014). Does optimal government size exist for developing economies? The case of Nigeria. *AAU Annals of Accounting, Educational & Social Research* 1(1), 15–33.

Amgain, J. (2017). Estimating optimal level of taxation for growth maximization in Asia. *Applied Economics and Finance*, 4(3), 47–55.

Andres, J. & Hernando, I. (1999). Does inflation harm economic growth? Evidence from the OECD. In *The Costs and Benefits of Price Stability*. University of Chicago Press, 315–348.

Armey, R. (1995). The Freedom Revolution. Washington D.C.: Regnery Publishing Co.

Asimakopoulos, S. & Karavias, Y. (2016). The impact of government size on economic growth: A threshold analysis. *Economics Letters*, 139, 65–68, https://doi.org/10.1016/j.econlet.2015.12.010.

Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98 (5), 103–125.

Barro, R. J. (1996). Determinants of economic growth: A cross-country empirical study. *National Bureau of Economic Research*, No. w5698.

Begum, H. & Aziz, Md. S. I. (2019). Impact of domestic credit to private sector on gross domestic product in Bangladesh. *IOSR Journal of Economics and Finance*, 10 (1), 45–54.

Bencivenga, V. R., Smith, B. D., & Starr, R. M. (1996). Liquidity of secondary capital markets: Allocative efficiency and the maturity composition of the capital stock. *Economic Theory*, *7* (1), 19–50.

Borensztein, E., De Gregorio, J., & Lee, J. (1998). How does foreign direct investment affect economic growth? *Journal of International Economics*, 45 (1), 115–135.

Bulutay, T. (1995). *Employment, Unemployment, and Wages in Turkey*. Ankara, International Labour Office State Institute of Statistics.

Chao, J. & Grubel, H. (1998). *Optimal Levels of Spending and Taxation in Canada*. The Fraser Institute, Vancouver.

Chobanov, D. & Mladenova, A. (2009). *What is the optimal size of government?* Working Paper 335309, Institute for Market Economics, Sofia, http://ime.bg/uploads/335309_OptimalSizeOfGovernment.pdf.

Christie, T. (2014). The effect of government spending on economic growth: Testing the non-linear hypothesis. *Bulletin of Economic Research*, 66 (2), 183–204, http://dx.doi.org/10.1111/j.1467-8586.2012.00438.x.

De Mendonça, H. F. & Cacicedo, T. (2015). Size of government and economic growth in the largest Latin American country. *Applied Economics Letters*, 22 (11), 904–910, https://doi.org/10.1080/13504851.2014.987909.

Devarajan, S., Swaroop, V., & Zou, H. (1996). The composition of public expenditure and economic growth. *Journal of Monetary Economics*, 37 (2), 313–44.

Di Liddo, G., Magazzino, C., & Porcelli, F. (2018). Government size, decentralization and growth: Empirical evidence from Italian regions. *Applied Economics*, 50 (25), 2777–2791, DOI: 10.1080/00036846.2017.1409417.

Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74 (366a), 427–431.

El Husseiny, I. A. (2019). The optimal size of government in Egypt: An empirical investigation. *The Journal of North African Studies*, 24 (2), 271–299, DOI:10.1080/13629387.2018.1428798.

Facchini, F. & Melki, M. (2013). Efficient government size: France in the 20th century. *European Journal of Political Economy*, 31, 1–14.

Forte, F. & Magazzino, C. (2011). Optimal size of government and economic growth in EU countries. *Journal of Analytical and Institutional Economics,* 28, 295–322.

Forte, F. & Magazzino, C. (2016). Government size and economic growth in Italy: A time-series analysis. *European Scientific Journal*, 12 (7) ISSN: 1857–7881 (Print) e - ISSN 1857- 7431 doi: 10.19044/esj. 2016.v12n7p149

Forte, F., & Magazzino, C. (2018). Wagner's law, government size and economic growth: An empirical test and theoretical explanations for Italy 1861–2008. In *Gustav von Schmoller and Adolph Wagner*, 129–151. Springer, Cham.

General Directorate of Budget and Fiscal Control, http://www.bumko.gov.tr/TR,4461/butce-gider-gelir gerceklesmeleri-1924-2016.html. (18 November 2018)

Gillman, M. & Harris, M. (2010). The effect of inflation on growth. *The Economics of Transition*, 18 (4), 697–714.

Grossman, P. (1987). The optimal size of government. *Public Choice,* 53, 131–147.

Günalp, B. & Dinçer, O. (2005). The optimal government size in transition countries. Department of Economics, Hacettepe University, Beytepe, Ankara, and Department of Commerce, Massey University, Auckland.

Gwartney, J., Lawson, R., & Holcombe, R. (1998). The size and functions of government and economic growth. *Joint Economic Committee*, Washington, D.C.

Herath, S. (2009). The size of the government and economic growth: An empirical study of Sri Lanka. *Institute for the Environment and Regional Development. SRE Discussion Working Paper 2009/08*, Vienna University of Economics and Business http://epub.wu.ac.at/2962/.

Hok, L., Jariyapan, P., Buddhawongsa, P., & Tansuchat, R. (2014). Optimal size of government spending: Empirical evidence from eight countries in Southeast Asia. *The Empirical Econometrics and Quantitative Economics Letters*, 3 (4), 31–44.

Jones, L. E., Manuelli, R. E., & Rossi, P. E. (1993). Optimal taxation in models of endogenous growth. *Journal of Political Economy*, 101, 485–519.

Kahveci, Ş. & Terzi, H. (2017). Türkiye'de doğrudan yabancı yatırımlar ve ekonomik büyüme arasındaki ilişkilerin nedensellik analizleri ile testi. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,* No: 49, 135–154.

Karras, G. (1996). The optimal government size: further international evidence on the productivity of government services. *Economic Inquiry,* 34, 193–203.

Karras, G. (1997). On the optimal government size in Europe: Theory and empirical evidence. *The Manchester School of Economic & Social Studies*, Blackwell Publishing, 65 (3), 280–294.

Kim, D. H., Wu, Y. C., & Lin, S. C. (2018). Heterogeneity in the effects of government size and governance on economic growth. *Economic Modelling, 68*, 205–216, https://doi.org/10.1016/j.econmod.2017.07.014.

King, R. & Rebelo, S. (1990). Public policy and economic growth: Developing neoclassical implications. *Journal of Political Economy*, 98, 126–151.

Kormendi, R. C. & Meguire, P. G. (1985). Macroeconomic determinants of growth: Cross-country evidence. *Journal of Monetary Economics*, 16 (2), 259–276.

Magazzino, C. (2014). Government size and economic growth in Italy: An empirical analysis based on new data (1861–2008). *International Journal of Empirical Finance,* 3(2), 38–54.

Makin, A., Pearce, J., & Ratnasiri, S. (2019). The optimal size of government in Australia. *Economic Analysis and Policy*, 62(C), 27–36.

Mallik, G. & Chowdhury, A. (2001). Inflation and economic growth: Evidence from four Asian countries. *Asia-Pacific Development Journal*, 8 (1), 123–135.

Mavrov, H. (2007). The size of government expenditure and the rate of economic growth in Bulgaria. *Economic Alternatives,* 1, 52–63.

Mittnik, S. & Neumann, T. (2003). Time-series evidence on the nonlinearity hypothesis for public spending. *Economic Inquiry,* 41, 565–573.

OECD. https://data.oecd.org/emp/labour-force-participation-rate.htm (18 November 2018)

Olowofeso, E. O., Adeleke, A. O., & Udoji, A. O. (2015). Impact of private sector credit on economic growth in Nigeria. *CBN Journal of Applied Statistics*, 6 (2).

Phillips, P. C. & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75 (2), 335–346.

Peden, E. (1991). Productivity in the United States and its relationship to government activity: An analysis of 57 years, 1929–1986. *Public Choice,* 69, 153–173.

Pesaran, H. M. & Shin, Y. (1999). Autoregressive distributed lag modeling approach to cointegration analysis. *In* S. Storm (Ed.) *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*. Chapter 11. Cambridge University Press.

Pesaran, M. H., Shin, Y., & Smith R. J. (2001). Bounds testing approaches to the analysis of level relationship. *Journal of Applied Econometrics*, 16 (3), 289–326.

Pevcin, P. (2004). Does optimal spending size of government exist? Paper presented at the European Group of Public Administration Conference, Ljubljana, Slovenia.

Presidency of Turkey, Strategy and Budget Office. http://www.sbb.gov.tr/ekonomik-veriler/#1540461995857-3570233a-09e6 (18 November 2018)

Rahn, R. & Fox, H. (1996). What is the Optimum Size of Government? Vernon K. Krieble Foundation.

Rezk, E. (2005). Public expenditure and optimal government size in an endogenous growth model: An analysis of the Argentine case. *The National University of La Plata*.

Saget, C. (2000). Can the level of employment be explained by GDP growth in transition countries? *Labour*, 14 (4), 623–643.

Sala-i-Martin, X. (2002). Sources of growth. *Macroeconomic Management: Programs and Policies* Ed. Mohsin S. Khan, Saleh M. Nsouli, Chang-Huey Wong, IMF Institute, Washington, DC., 58–97.

Scully, G. (1994). What is the optimal size of government in the US? *National Center for Policy Analysis*, Policy Report No. 188.

Scully, G. W. (1998). Measuring the burden of high taxes. *National Centre for Policy Analysis–Policy Report*, No: 215, Mimeo.

Scully, G. W. (2003). Optimal taxation, economic growth, and income inequality. *Public Choice*, 115 (3), 299–312.

Scully, G. (2008). Optimal taxation, economic growth and income inequality in the United States. *National Center for Policy Analysis*, Policy Report No. 316.

Şen, H. & Kaya, A. (2019). Alternatif göstergeler bazında Türkiye'de optimal kamu kesimi büyüklüğünün tahmini. *Bankacılar Dergisi*, 109, 49–81.

The Ministry of Treasury and Finance. Annually Budget Justifications.

The Ministry of Treasury and Finance. Bütçe Gider ve Gelir Gerçekleşmeleri (1924–1995). A statistical report, revised 2nd Edition.

Vedder, R. & Gallaway, L. (1998). Government size and economic growth. *Joint Economic Committee*, Washington D, 1–15.

Wooldridge, J. (2012). Panel data models with heterogeneity and endogeneity. *Institute for Fiscal Studies*.

World Bank. www.worldbank.org. (1 November 2018)

Yam, T. K., Wu, F., Heng, T. M., Ping, J. T., & Sia, H. Y. (2002). Employment growth will trail economic recovery in 2002: Evidence from econometric estimates. *Economic Survey of Singapore*, First Quarter, 45–52.

Received: March 14, 2021
Accepted: February 28, 2022

APPENDIX

Table A1: Descriptive statistics

	RGdpG	CE	IE	CIE	DE	EE	HE	EHE	ECenGov	RCenGov	RDirT	RIndT	INF	FDI	DCPS	EMP
Mean	4.54	11.80	2.59	14.40	1.99	2.75	0.78	3.54	19.11	16.19	5.39	7.68	40.95	0.80	25.73	54.99
Maximum	11.11	15.77	4.40	18.91	3.01	4.60	1.60	5.80	33.54	24.28	7.22	12.77	105.21	3.65	69.85	67.50
Minimum	−5.96	7.52	0.98	9.70	1.20	1.39	0.31	1.72	11.88	9.65	3.86	3.34	6.25	0.01	13.58	46.20
Std. Dev.	4.30	2.07	0.78	2.38	0.52	0.88	0.38	1.23	5.65	4.44	0.76	3.45	29.48	0.89	15.64	6.42
Skewness	−0.88	−0.49	−0.14	−0.11	0.01	0.31	0.63	0.30	0.66	0.12	0.22	0.28	0.38	1.53	1.72	0.47
Kurtosis	3.10	2.67	2.30	2.33	1.84	2.14	2.10	1.90	2.79	1.61	2.52	1.37	1.92	4.72	4.62	2.13
Jarque-Bera	5.58	1.96	1.03	0.89	2.38	2.04	4.31	2.83	3.22	3.53	0.75	5.29	3.11	22.14	26.01	2.94
Probability	0.06	0.37	0.59	0.63	0.30	0.35	0.11	0.24	0.19	0.17	0.68	0.07	0.21	0.00	0.00	0.22

Source: Author's computation

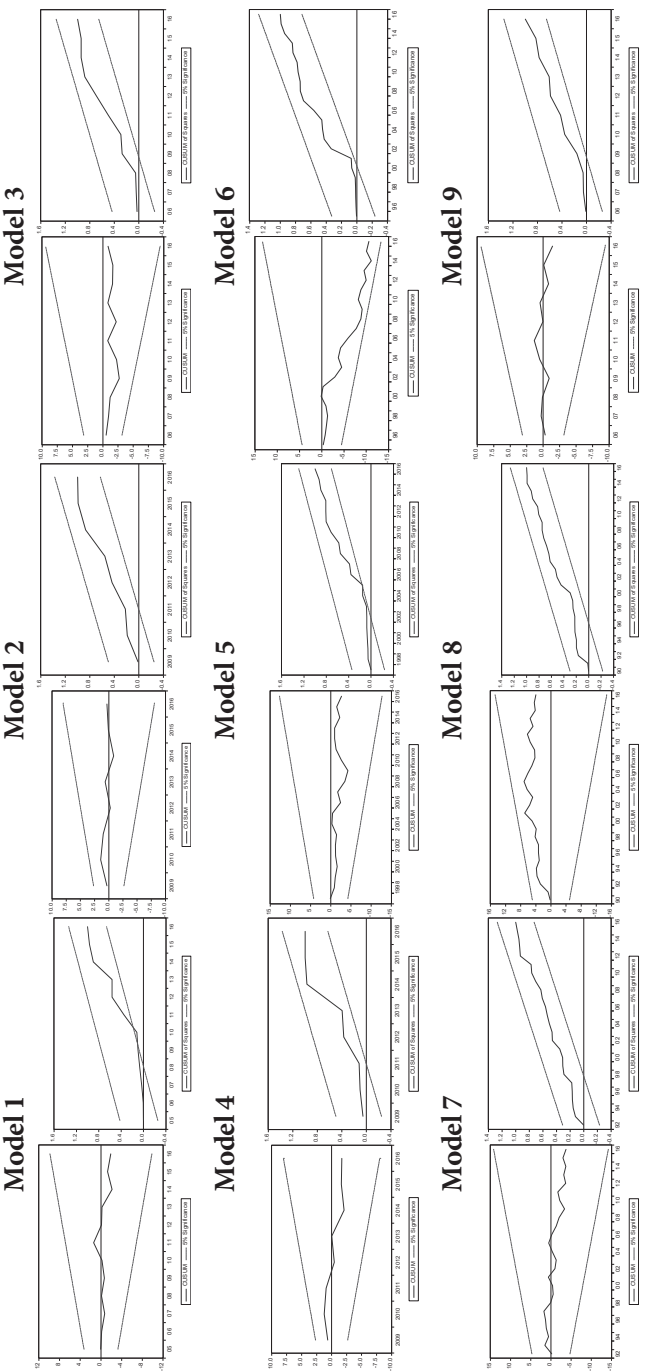


Figure A1: CUSUM and CUSUM-Q test results

ECONOMIC ANNALS, Volume LXVII, No. 232 / January – March 2022
UDC: 3.33 ISSN: 0013-3264

*Sonja Avlijaš**

BOOK REVIEW: The Post-Crisis Developmental State. Perspectives from the Global Periphery, by Tamas Gerocs and Judit Ricz, Cham, Switzerland: Palgrave Macmillan, pp. 337.

The global financial and economic crisis of 2008–2009, and more recently the COVID-19 pandemic, have highlighted the growing challenges to global capitalism. After almost half a century of increasing market and trade liberalisation, the state is back in vogue throughout the world. This volume, edited by Tamas Gerocs and Judit Ricz, starts from the well-established empirical fact that developing and emerging market economies, which they refer to as the global semi-periphery, differ from the more advanced economies regarding not only their economic and social structure but also their capacity and autonomy to implement the developmental policies that are now being promoted in many contexts. They invite us to consider what the newly emerging post-crisis paradigm of greater state involvement in the economy means for states which are institutionally, financially, and/or politically under-capacitated, and where domestic agency and policymaking are constrained by exogenous power structures such as multinational corporations and the geostrategic interests of large countries.

This is an urgent, albeit very complex, question for both academia and policymakers. The volume adds much-needed nuance to the habitual dichotomy of the state vs. the market that often permeates political discourse. By adopting a comparative analytical framework that accounts for the varieties of state capitalism and by considering international as well as domestic political economy, the volume enriches the discussion in economics on the developmental state. It also contributes to the field of political economy itself by introducing the view

* University of Belgrade – Faculty of Economics, email: sonja.avlijas@ekof.bg.ac.rs, ORCID: http://orcid.org/0000-0003-3270-0039

from the global semi-periphery, where the interaction between international factors and domestic institutional weaknesses and political instability often has strong effects.

In her introductory chapter and in chapter 5 of the volume, Judit Ricz recognises a post-crisis developmental opportunity for the semi-periphery, as the structure of the global economy changes towards decentralisation of productive relations and an increasingly important role for global value chains that include smaller, dispersed economies from around the world. However, she also highlights the paradox of the growing role of populist, ethnonationalist, and patriotic elites that use economic nationalism to justify their autocratic tendencies while benefiting economically from globalisation. These puzzling trends have increased the global demand for a better understanding of the semi-peripheral economies that are plugging into the global value chains, making this volume a very timely read.

Ricz' introduction sets the stage for the next three chapters, which provide an overview of development and change in the era of modernisation. In Chapter 2, Laszlo Csaba reminds us of the strong influence institutional economics has had on policymakers during the 'liberal' era. He rejects the state vs. market dichotomy and argues that we need to improve our understanding of the state institutions that underpin and coordinate market economies. He emphasises the important distinction between being statist and being pro-institution-building and warns of the phenomenon of institutional hollowing-out that is occurring in many semi-peripheral contexts such as Hungary and Poland, and to a much stronger degree in Russia and Ukraine, where institutions that are supposed to support the market economy exist formally but have been de facto crippled by political means. He concludes that we need to pay particular attention to informal institutions such as the sets of values that prop up the market system and argues that we are in the dark when it comes to understanding how shared norms and values work in this context and why social trust and efficient institutions work in some places but not in others.

Csaba broaches a very important subject here, which could be probed further. While the rise of global capitalism has been associated with an unprecedented rise in inequality and growing delegitimisation of institutions even in the west, institutional economists have not paid much attention to the exploitative and unequalising effects of the formal and informal institutions that coordinate markets. The assumption is that 'properly' implemented institutions lead to superior developmental outcomes, while the reasons behind their growing popular delegitimisation are not explored or these grievances are simply

dismissed as populism. However, research on how market coordination shapes inequalities is beginning to emerge in comparative political economy, and I hope that eventually it will also permeate political economy research on the global semi-periphery.

In Chapter 3 Andrei Yakovlev discusses the role of information and communication technologies (ICT) in the next stage of human development in the context of growing global interdependence between countries. He makes the valuable point that both dependent market economies and the 'Varieties of Capitalism' models that are state led are reactionary since they evolved as a response to globalisation, i.e., they were shaped by it. By contrasting this to the western coordinated and liberal market economies which developed their institutional make-up prior to widespread liberalisation, he is perhaps suggesting that the latter have had more leverage to resist the adverse effects of market liberalisation, although he does not state this explicitly. Bringing global politics into the picture, he recognises the role of the elites in big emerging countries which have benefited from globalisation but remain unsatisfied with their role in the world order and argues that this is why they are driving the global push towards state-led developmentalism. He ends by exploring what it would take for different elites to cooperate within countries and for countries to cooperate in today's system of high economic interdependence. He sees this cooperation as a new cycle of development following ideological conflict and competing economic systems and argues that in this switch from competition to cooperation in the global arena the personalities of leaders are key.

While I undoubtedly root for global cooperation, I am not clear how it would work in the context of the ICT-led era of growth. The global economic system has been switching away from economies of scale and towards seizure of market shares and intensification of winner-takes-most dynamics of digitalisation. We have even struggled to share medical supplies and vaccines during a major global pandemic. The world is also facing a major fight over the raw materials and energy sources that are needed for continued growth and digitalisation, while we are also attempting to transition to 'green' growth. Thus, while I agree that in current circumstances competition can lead the world to a race to the bottom and that cooperation is the only solution, I would like to know how this cooperation is envisaged in a world where global elites are de facto fighting over resources, and where inequalities keep growing. What would be the basis of this global cooperation, and what would replace the market-grabbing tendencies of multiple global players that we have seen so far in the era of ICT-driven globalisation? Can we really afford to go on assuming that growth has trickle-down effects and that

continuing business as usual, but with some vague idea of cooperation between elites, will resolve the issue of inequalities and also the climate crisis, which is closely connected to the resource-acquisition-oriented conflicts that we see unfolding? It is very difficult for me to imagine a world where everything stays the same in terms of the economic system that we currently have, yet we somehow miraculously start cooperating on a global scale by way of elite leadership. I would thus be interested to hear more about the systemic conditions under which this form of global cooperation is envisaged.

In Chapter 4, Miklos Szanyi poses the interesting question of whether catching-up countries benefit more from periods of accelerating technological development or from the widespread deployment phase of cutting-edge technologies. He concludes that the successful catching-up countries have been better at utilising the business opportunities of the deployment phase and suggests that capital and skilled labour are necessary to seize these opportunities. I would also add insights from economic geography on the importance of networks and politics in how places utilise the business opportunities that present themselves, which fits nicely with Csaba's chapter 2 on our need to better understand the informal institutions that structure developmental outcomes.

The second part of the book starts with Judit Ricz' overview of the literature on how developmental states have emerged, from the political economy perspective. She juxtaposes two schools of thought: the role of elites as agents of development vs. the urgency theory where development is the result of coalitional and social demands pressurising political elites. She then moves on to the historical and cultural factors that have shaped East Asian developmental states, including homogenous societies and good quality institutions. She points out the role of strong leaders in East Asian developmental states and explains that although broad swaths of the population were often politically excluded and sometimes even repressed, they were never ignored economically, so the idea of equitable distribution of gains from growth was arguably part of the growth model and economic inclusion of the broader society mattered. However, where the institutions and policies that lead to long-term sustainable development come from remains an unanswered question. What might help us answer this question, Ricz argues, is a renewed focus on domestic agency and power dynamics not only among elites but also among broader sections of the population in the semi-periphery. It would be interesting to learn more about the role of authoritarianism in developmental states, and to what extent this framework is applicable in the context of rising authoritarianism in Eastern Europe.

Case studies of Slovakia, Ethiopia, Namibia, and Zimbabwe in the subsequent chapters offer empirical input on the role of domestic agency. Cases from sub-Saharan Africa provide especially insightful accounts on how domestic political factors interact with exogenous constraints and the international political economy of globalisation. The Slovakian case is unique in the volume in that it points to the unequalising effects of developmental policies, with growing disparities between the urban core around the capital city and the rest of the country. The fact that Slovakia is referred to as a success story in the concluding chapter is somewhat confusing, because the general message of the volume seems to be that broader swathes of the population need to benefit from developmental policies in order for them to work. However, this message is not explicit, so I am left somewhat in the dark as to whether the authors assume that broader development gains are the result of trickle-down growth, or whether some populations are expected to, and often do, take up the disproportionate cost that developmental policies impose onto them. Some chapters suggest that non-economic policy areas such as education matter, while others focus on the idea that political stability is a key pre-condition, one which often remains unfulfilled due to both domestic and international constraints. There seems to be some tautology in this line of argument, because we are left with the idea that local political conflicts need to be resolved in order for countries to develop institutions and spur growth, but at the same time the fact that it is mostly the elites that are benefiting from growth in today's world is also a factor of political destabilisation. Moreover, we know that some policies are good for development, but they cannot be implemented because no political consensus can be formed around them, and we do not understand why.

The last part of the book brings the developmental state perspective into a conflictual relationship with trade policy, thus adding an additional layer of international political economy analysis which complicates the picture and has important implications for the global economy as a whole. In Chapter 13, in the context of North Africa, Tamas Szigetvari discusses the role of elite corruption and the adaptation of free trade institutions to particular interests, further emphasising the point made by Laszlo Csaba in Chapter 2 that statism is not the same as state coordination of economic policy and market relations, and that we need to be careful whether we are actually advocating institutional regulation or whether we are strengthening the role of elites by providing them with a theoretical and policy framework that can serve to hollow out the institutional foundations of global capitalism. The cases of Brazil and Pakistan appear to push the argument in the opposite direction, showing that certain international

processes can have a destabilising effect on local political dynamics and even developmental policies, and thus threaten their effectiveness.

After reading this volume, I am even more convinced that the post-crisis developmental paradigm is a major challenge of our times. Almost none of the country cases covered in the volume are cases of successful developmental states (the exceptions are Taiwan, and partially Ethiopia and Slovakia), so the volume in fact provides an overview of developmental state failures that need to be tackled in the process of coming up with a developmental paradigm for the global semi-periphery.

I am now also even more convinced that within-country inequalities need to be high on the agenda of the new developmental state paradigm, and that it is no longer enough to assume that growth-with-equity means that developing countries catch up with advanced economies 'on average'. Distribution of development gains within countries has to be high on the catching-up agenda. As long as we are not willing to find ways to redistribute resources more equitably, rising inequalities within nation states will decrease our chances of finding political solutions to current challenges.

This volume is a welcome contribution to the new developmental state paradigm that is still in the making. It raises important questions, but comes short of unpacking the catching-up/growth agenda for the semi-periphery and its feasibility in the context of rapidly growing within-country inequalities and the global climate crisis. It is encouraging that issues of political economy have finally entered the economic arena. I look forward to continuing this debate in a direction which integrates the issue of semi-periphery catch-up with the growth paradigm crisis that we are seeing in advanced capitalist economies that have 'all the right ingredients' in terms of their institutional make-up.

Received: February 28, 2022
Accepted: March 17, 2022

INSTRUCTIONS TO AUTHORS

Economic Annals is an international professional journal published quarterly by the Faculty of Economics, University of Belgrade. It publishes research in all areas of economics and business administration, particularly for transition and emerging economies. The journal encourages the submission of original unpublished works, not under consideration by other journals or publications. Contributions written in English and in electronic form should be forwarded to: ea@ekof.bg.ac.rs.

The journal maintains high scientific standards. Papers submitted for publication should be original, relevant and scientifically accurate. Authors are expected to provide new information or analysis and should present a summary of the conclusions they draw, maintaining coherence and compactness of their reasoning. The originality of the work is subject to test by iThenticate crosscheck. The texts should also follow appropriate technical standards and stylistic criteria. UK spelling (specialisation, labour, etc.) should be used, while both UK and US abbreviations are acceptable.

All submitted papers will undergo a double-blind refereeing process. An anonymous version of the paper should be submitted along with a separate cover page, containing the article's title, author's name and affiliation, e-mail address, and a suggested running head (an abbreviated form of the title of no more than 50 characters with spaces). The cover page should also should contain a short abstract of between 100 to 200 words, summarising the major points and conclusions of the paper, a list of up to five keywords and up to five two-digit codes in accordance with the Journal of Economic Literature (JEL) classification (https://www.aeaweb.org/econlit/jelCodes.php).

Papers should be prepared as a single file (including text, notes, references, figures, and tables) in MS-Word. Papers should be submitted in A4 page format (21×29.7 cm), left and right margins cm, top and bottom 4.5 cm, in font Times New Roman 12, single line spacing. There should be no tabs in paragraphs and a single line spacing should separate paragraphs and titles. Tables, figures and footnotes should be included as they are intended to appear in the final version. Footnotes (in font 10) should be kept to a minimum and numbered as superscripts. Papers which do not conform with the above instructions will not be taken into consideration.

As a rule, submitted articles should not exceed 8,000 words. All pages apart from the first one should be numbered. Subtitles should be concise, clearly marked in bold, and numbered (up to two levels of numbering). No other entries should be bolded. Formulae should be numbered on the right-hand side of the page. In case of long proofs, these should be inserted in a separate Appendix, following the References. Tables and Figures must not use colour, and should be in a format easy to edit, for instance they should take half a page (or a full page) within the indicated margins. They should be clearly labelled at the top, with a legend at the bottom, and should be logically ordered, using Arabic numerals. Sources of the data should be given below tables and figures.

Papers should follow APA style guidelines: https://apastyle.apa.org/style-grammar-guidelines/references/examples#textual-works. Some key points watch out for are as follows. Parenthetic references in the text and in footnotes should be listed by the author surname, with the year of publication in parentheses; in case of more than one author use an ampersand, for instance: (Atkinson, Picketty & Emmanuel, 2011). Narrative citations within the text should use "and" rather than ampersand, for instance: Djankov, Glaeser and La Porta (2003). Use an ampersand in the list of references. When citing works with one or two authors, include the author name(s) in every citation. For works with three or more authors, include the name of only the first author plus "et al." in every citation (even the first citation). Include all author names in the list of references. If the author is unknown, the first few words of the reference should be used; this is usually the title of the source. For example: (*A guide for economy*, 2019). Multiple works by the same author are sorted by date in ascending order; if the works are in the same year they should be ordered alphabetically by title and allocated a letter (a, b, c, ...) after the date. Cite only the works you have read and incorporated into your writing; readers may find long strings of citations difficult to understand. Before submitting your paper, check that all references cited in the paper are included in the reference list at the end of the paper, and that all papers included in the reference list have been cited in the text.

References should be left aligned in alphabetical order in the reference list, according to the following formats:

Article in journals

Author surname(s), initial(s). (Year). Article title. Journal, Volume number (issue or part number, optional), page numbers. DOI.

Rodrik, R., Subramanian, D., & Trebbi, F. (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, 9(2), 131-165.

https://DOI: 10.1023/B:JOEG.0000031425.72248.85.

Book

Author surname, initial(s). (Year). *Title*. Publisher location: Publisher

De Grauwe, P. (2020) *Economics of Monetary Union* (13th ed.). Oxford: Oxford University Press.

Edited book

Author surname, initial(s). (Ed(s).). (Year). *Title*. Publisher location: Publisher

Baltagi, B.H. (Ed.). (2003). *A Companion to Theoretical Econometrics*. Oxford: Blackwell

Book with several authors

When there are multiple authors, list them all, with the addition of ampersand (&) before the last surname. If there are more than seven authors, list the first six, then write three full stops (…), and at the end write the last author.

Acemoglu, D., & Robinson, J.A. (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge: Cambridge University Press.

Baumol, W. J., Panzar, J. C., & Willig, R.W. (1982). *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt, Brace, Jovanovich, Inc.

Chapter in book

Last name of the chapter author, initial(s). (Year). Chapter title. In editor initial(s), surname (Ed.). *Title* (ed., pp.). Publisher location: Publisher

McMillan J., & Woodruff C. (2003) The central role of entrepreneurs in transition economies. In G. S. Fields, & G. Pfefferman (Eds.). *Pathways Out of Poverty* (pp. 105-121). Dordrecht: Springer. https://doi.org/10.1007/978-94-010-0009-3_6.

E-Book

Author surname, initial(s). (Year). *Title*. URL

Perry, R.B. (1909). *The Moral Economy.*

https://manybooks.net/book/137844/read#epubcfi(/6/2[id00000]!/4/2[id00000]/1:0)

Technical Reports or Working Papers

Individual authors

Author surname, initial(s) or corporate name. (Year). Title. (Report or Working Paper No.). URL.

Cătuți, M., Kustova, I. and Egenhofer, C. (2020) *Delivering the European Green Deal for Southeast Europe: Do we need a regional approach?* (CEPS Research Report No. 2020/1). https://www.ceps.eu/wp-content/uploads/2020/06/RR_2020-01_European-Green-Deal-for-South-Eastern-Europe.pdf.

Corporate authors

American Psychological Association, Task Force on the Interface Between Psychology and Global Climate Change. (2009). *Report of the APA Task Force on the Interface Between Psychology and Global Climate Change*.

http://www.apa.org/science/about/publications/climate-change.aspx

Newspaper Articles

Author surname, initial(s). (Year, Month Day). Title. *Title of Newspaper*, p. or pp. URL*

*only include if the article is online.

Note: the date includes the year, month and date.

Smialek, J. (2020, May 2). Hotel Group Will Return Tens of Millions in Small Business Loans. *The New York Times*, pp. 10.

https://www.nytimes.com/2020/05/02/business/economy/ashford-hotels-virus-monty- bennett.html

Website

Author surname, initial(s). (Year, month day). *Title*. URL

Mitchell, J.A. (2017, May 21). *How and when to reference*.

https://www.howandwhentoreference.com