Access to Justice and Economic Development: Evidence from an International Panel Dataset

A. Deseau, A. Levai, and M. Schmiegelow

Discussion Paper 2019-9

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

IRES

UCLouvain

Access to Justice and Economic Development: Evidence from an International Panel Dataset *

Arnaud Deseau[+], Adam Levai[++], and Michèle Schmiegelow[+++]

[+]Université Saint-Louis – Bruxelles, CEREC
[++]UCLouvain, CRIDES, IRES
[+++]UCLouvain, CRIDES

April 2019

Abstract

We empirically investigate the impact of access to justice (ATJ) on GDP per capita growth in a panel of 83 countries from 1970 to 2014. Our analysis relies on a new database documenting the number of judges per capita as a proxy for capturing the cross-country evolution of ATJ. The proxy measures the extent to which disputes between economic actors can be resolved at a relatively low cost, without dysfunctional delay and discrimination. In a dynamic panel setting using internal instruments, we find that increasing ATJ by 1% increases the five-year growth rate of GDP per capita by 0.86 p.p. (0.17 p.p. annually) with diminishing marginal returns. In line with the diminishing marginal returns argument, we find that the effect of ATJ is two times smaller in Europe compared to other regions due to higher levels of ATJ. We find no evidence of a differential effect of ATJ across other regions, income levels, legal origins, democracy, corruption of the judicial system or human capital levels.

Keywords: Access to Justice, Legal development, Economic Development, Growth, Institutions

JEL Codes: E02, K00, O11, O43, O47

*We would like to thank Alexia Autenne, Frédéric Docquier, Filip Dorssemont, Alexandre Girard, William Parienté and Henrik Schmiegelow for their invaluable comments. We also thank the participants at seminars and workshops at University of Bonn, UCLouvain, University Saint-Louis Bruxelles, Erasmus University Rotterdam and Spring Meeting of Young Economists for helpful comments and suggestions. We are grateful to: the Fonds Quetelet (Brussels) for granting us access to its cross-country collection of statistical yearbooks; the Ministries of Justice in Botswana, Belgium, Croatia, Estonia, Finland, Ireland, Latvia, Mali, Malta, New Zealand, Niger, Poland, Romania, Slovenia, South Korea, Sweden, Uruguay, Venezuela; and to National Statistical Service of Australia, Armenia, Qatar and United States for sharing their official data with us. This work was supported by the Fonds de la Recherche Scientifique - Projet de Recherches (FNRS-PDR), under the grant number "26036434". The usual disclaimers apply.

"*Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels*"

Sustainable Development Goal 16 - United Nations (2015)

1 Introduction

A country can have a very well-defined legal system and legal rules but if they are not accessible for the majority of people, due for example to high costs and long (dysfunctional) delays, then the legal system and rules, both substantive and procedural, will only have a limited impact on economic activity. For example, even though France and Senegal (a former French colony) have very similar law codes, the effective access to justice (ATJ) of citizens in both states is significantly different, leading thus to different economic outcomes.

The legal origin theory shows how differences in legal codes may affect financial development (La Porta et al.; 1997, 1998, 1999; Djankov et al.; 2007), labor markets (Botero et al.; 2004), competition (Djankov et al.; 2002, 2003) and finally economic growth (Beck et al.; 2000; Levine et al.; 2000; Mahoney; 2001).[1] More recent research emphasizes the role of justice quality and efficiency in fostering entrepreneurship, credit, agricultural and industrial activities at the country level (Chemin; 2009a,b; Jappelli et al.; 2005; Visaria; 2009; Amirapu; 2017). Yet to date, surprisingly little is known about the importance of ATJ in explaining cross-country income per capita differences.

As a first step in bridging this gap in the literature, we build a new database documenting the number of judges from 1970 to 2014 for 107 countries worldwide. This measure serves as a macro level proxy for ATJ, capturing the extent to which disputes can be resolved at a relatively low cost, without dysfunctional delays and discrimination. Using the panel structure of our data, we are able to address endogeneity issues and look at the impact of

[1]See La Porta et al. (2008) for a complete review of the contribution of the legal origin theory to the literature.

ATJ on economic growth using a difference GMM estimation.[2] Our benchmark specification suggests that a 1% increase in ATJ leads to, on average, a 0.86 p.p. increase in the five-year GDP per capita growth rate with diminishing marginal returns. This finding is robust to different subsamples of countries.

We then assess whether ATJ can foster economic growth differently across continents/regions, income groups, legal origins, political regime, corruption of the judicial system and human capital levels. Our results show no statistical difference in the effect along those dimensions. Only in the case of Europe, we find a significantly weaker effect: the impact of ATJ is almost two times lower compared to the rest of the world, which is in line with the diminishing marginal returns argument.

This paper makes three main contributions. First, we add to the empirical law and economics literature by providing a new database tracking the number of judges per capita in 107 countries between 1970 and 2014. This is notable as there is very few panel data allowing for cross-country comparisons of the effective judicial supply both for developed and developing countries. Moreover, it is the first database offering a comprehensive proxy for ATJ across country and through time. This is crucial as the United Nations has currently no historical data which would allow them to track the progress of the Sustainable Development Goal 16 from the ATJ aspect.

Second, we add to the literature analyzing empirically the determinants of economic growth as, to the best of our knowledge, we provide the first panel cross-country analysis looking at the effect of ATJ on economic growth. On that aspect, we complement the scant literature trying to explain cross-country income and growth differences by difference in the judicial characteristics (Berkowitz et al.; 2003; Feld and Voigt; 2003; Voigt et al.; 2015).

Third, we advance the existing literature on the macroeconomic effect of justice by addressing endogeneity issues in a more satisfactory way than existing contributions. While

[2] Our specifications focus on the effect of access to justice in log (as proxied by the number of judges per 100.000 inhabitants) on growth in GDP per capita. For brevity we will sometimes describe this as, with some abuse of terminology, "the impact of ATJ on growth" or "the impact of access to justice on economic development" (rather than the impact of access to justice on GDP per capita growth).

endogeneity is a long recognized issue in cross-country analysis, recent literature has advanced in better panel cross-country identification strategies. Benefiting from the panel structure of our data, we are able to use a difference GMM estimation which relies on first difference, to mitigate the confounding effect of country-level unobserved heterogeneity, and internal instrumentation to address reverse causality between ATJ and economic growth.

The rest of the paper is organized as follows. Section 2 discusses the relevance of our proxy and presents our new database with some stylized facts. Section 3 explains the econometric specification and estimation methods. Section 4 presents the results of the aggregate effect of access to justice on economic growth, evaluates its robustness, and explores the heterogeneity of the impact regarding income level, region, legal origin, political regime, judicial corruption and human capital. Section 5 concludes.

2 Access to Justice

2.1 Theory and Measurement

ATJ is a difficult concept to define, as it comprises many dimensions which are not solely economic. We can decompose ATJ into four main dimensions: 1) information, 2) social, 3) geography and 4) cost dimension (Woolf; 1996; Hammergren; 2014). A society with a high level of ATJ should enable the majority of its citizens to: know their rights, provide them with information on which steps to undertake, and which competent people to meet in order to introduce a plaint or defend themselves against it (dimension 1); they should not face discrimination of any kind along the judicial process (dimension 2); they should be capable to go effectively to court, meet the legal staff and exchange with them (dimension 3); finally, they should be facing a procedural cost which is not prohibitive both in terms of monetary and time cost (dimension 4).

The complexity of ATJ makes it difficult to measure accurately across countries and through time. For this reason, there are very few publicly available datasets that one can exploit for a cross-country comparison of ATJ. In addition, for existing datasets, the time dimension is small which is not suitable for a meaningful panel analysis. One example is *The World Justice Project*: this non-profit organization has collected an impressive amount of data via general population polls and questionnaires sent to in-country law professionals to build an index describing the *Rule of Law* at the national level. One of the 44 sub-factors of the Rule of Law Index concerns directly ATJ as it measures if "people can access and afford civil justice" (sub-factor 7.1) but is available only from 2012. Another initiative launched by *The World Justice Project* called *Global Insights on Access to Justice* have gathered data on legal needs and public ATJ in 45 countries but again, the time dimension is small as the first available year is 2017.

We argue that the total number of all types of professional judges per capita is a relevant proxy for ATJ at the macro level. We can identify two main channels through which the density of judges directly affects ATJ. First, the number of judges per capita directly improves ATJ through a quantity channel: more judges means more resolved cases, reducing both the monetary and time cost of accessing to justice (dimension 4). Specifically, the literature on the determinants of court output has established that this channel is more important for developing, middle-income economies and small-size courts (Dimitrova-Grajzl et al.; 2016; Grajzl and Silwal; 2017) rather than for developed countries or urban courts (Beenstock and Haitovsky; 2004; Dimitrova-Grajzl et al.; 2012).

Second, the literature provides evidence that the number of judges per capita can affect ATJ through the quality of justice channel. At first glance, one can argue that the positive effect of the quantity channel can be cancelled out by a decrease in the quality of justice. However, most of the empirical studies reject the existence of a quantity-quality trade-off within courts (Rosales-López; 2008; Coviello et al.; 2015; Dimitrova-Grajzl et al.; 2016; Grajzl and Silwal; 2017). Moreover, in a cross-section of 36 European countries, Voigt and El-Bialy (2016) shows that more judges per court is even positively associated with a measure of judicial independence (quality of justice) while it is not significantly correlated with the resolution rate (quantity of justice). This last result provides indications that adding more judges in an already developed judicial system can improve the quality of justice. A greater quality of justice can be measured by reducing the number of appealed cases or verdict reversals.

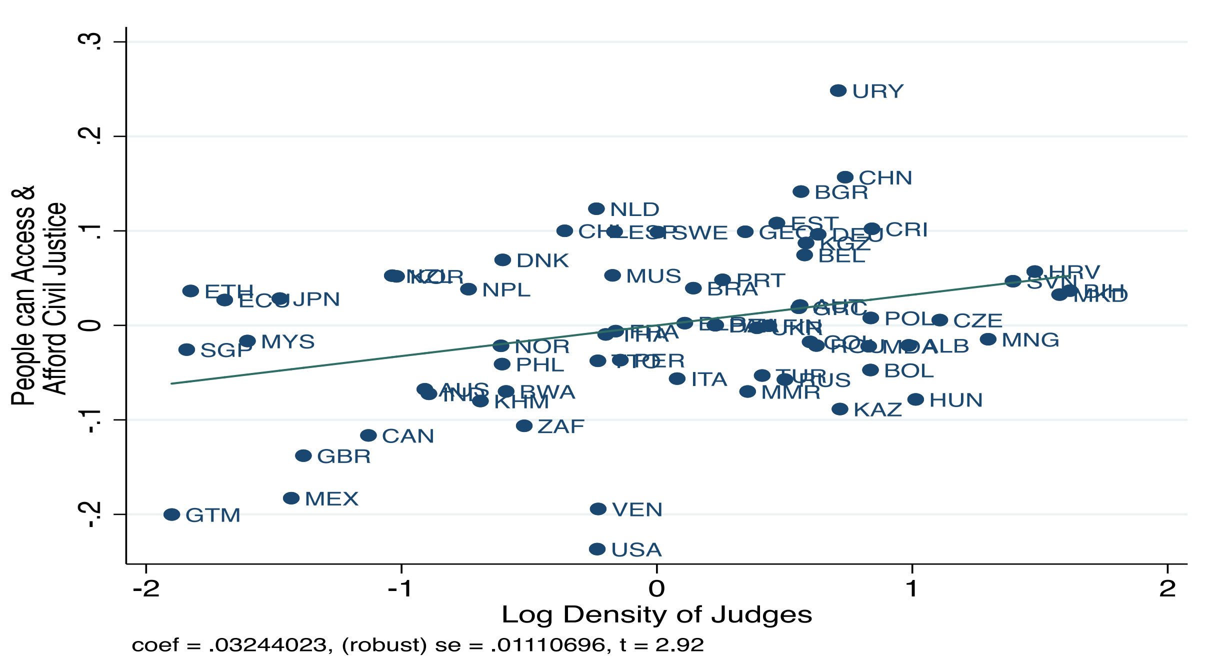
Overall, we view our proxy as capturing shocks to the aggregate supply of justice: an increase in the number of judges per capita decreases the trial length and increases the total number of treated cases, ceteris paribus. Consequently, ATJ is enhanced primarily via a reduction in both monetary and time cost of justice.

Our measure is coherent and transparent as the data is collected using official sources. Data on the number of judges are gathered from public institutions, international organizations and academic publications (see table A3 for more details) and are, for the most, publicly available. Moreover, the judge has the universal role of supplying justice by resolving disputes in a court of law across all national legal systems, allowing for cross-country comparisons.

Figure 1 provides an empirical validation of our proxy by showing a strong and positive correlation between our measure of ATJ (log density of judges) and the access to civil justice score as measured by the 2019 edition of *The World Justice Project - Rule of Law Index*, conditional on log of GDP per capita. The log density of judges is averaged across the whole 2000-2014 period to maximize sample size. In particular, the score constructed by the World Justice Project measures: "the accessibility and affordability of civil courts, including whether people are aware of available remedies [dimension 1]; can access and afford legal advice and representation [dimensions 3 and 4]; and can access the court system without incurring unreasonable fees [dimension 4], encountering unreasonable procedural hurdles, or experiencing physical or linguistic barriers [dimensions 2 and 3]". Figure 1 therefore gives direct evidence that the density of judges is a good proxy for ATJ as it captures all four dimensions described above.[3]

[3] In appendix figure A1, we test the quality of our proxy to alternative measures depicting more the dimension 2, 3 or 4 of ATJ with similar results.

Figure 1: The Partial Effect of the Density of Judges on Access to Civil Justice



Note: This figure plots the relationship between the log density of judges (averaged between 2000-2014) and the access to civil justice score of *The World Justice Project* (year 2019). This adjusted partial residual plot is based on an OLS regression where the log GDP per capita is used as a control variable and with robust standard errors.

Obviously, our proxy comes with some disadvantages. First, we measure access to a formal court of justice, neglecting the effect of informal justice and out-of-court settlements despite their importance in some developed and developing countries (Galanter; 1981; Platteau; 2015). This is a concern as we tend to underestimate the ATJ in some countries. However, many alternative dispute resolution mechanisms such as mediation require court-backed enforcement and thus the intervention of formal justice. In addition, in a context of competition between informal and formal legal institutions (known as legal pluralism), Aldashev et al. (2012) shows that formal law acts as an outside anchor on the custom. In that perspective, increasing access to the formal justice improves outside options of the plaintiff, letting him the choice to exit the informal system or forcing custom to adjust.

Second, due to data availability, we collect data on professional judges in all types of courts: first instance, second instance (appellate courts), supreme courts etc.; dealing with all types of cases: civil, criminal or other types of cases. A priori, we cannot say that some type of judges dealing with some type of cases are not important for economic activities. Judges often act as generalists: they adjudicate all relevant cases both civil and criminal. Only in large courts and developed economies judges tend to specialize in one specific domain (e.g., marital matters for civil cases). The civil cases are potentially the most important type of cases for economic development since they deal directly with economic activity and they have the highest monetary value for an average case. Moreover, the majority of legal issues are civil rather than criminal (Pleasence et al.; 2013). Therefore, by considering professional judges dealing with all types of cases can only bias our results downwards since an additional professional judge dealing with civil cases (compared to a judges dealing with criminal cases) contributes more to economic development. Figure 1 indicates that we are well capturing access to civil justice.

Finally, in our empirical exercise we focus on the quantity rather than quality of judges, while the quality of judges can influence ATJ via the same channels of quantity and quality of judgments (Ramseyer; 2012; Bielen et al.; 2018). That omission is a concern only in the case of a specific trend in the quality of judges on the period 1970-2014. Although the quality of judges might have changed over time, we think that this evolution stays minor compare to the development of the number of judges on the same period. Moreover, we control for time and country fixed-effects in our specifications, capturing the general trend and country specificities in terms of education which are good approximation of the quality of judges.

2.2 Data and Stylized Facts

Figures on the number of judges used to proxy ATJ have been gathered firstly from international organizations such as the Inter-university Consortium for Political and Social Research (ICPSR), the Commission Européenne Pour l'Efficacité de la Justice (CEPEJ), the United Nations Office on Drugs and Crime (UNODC) and the Organization of American States

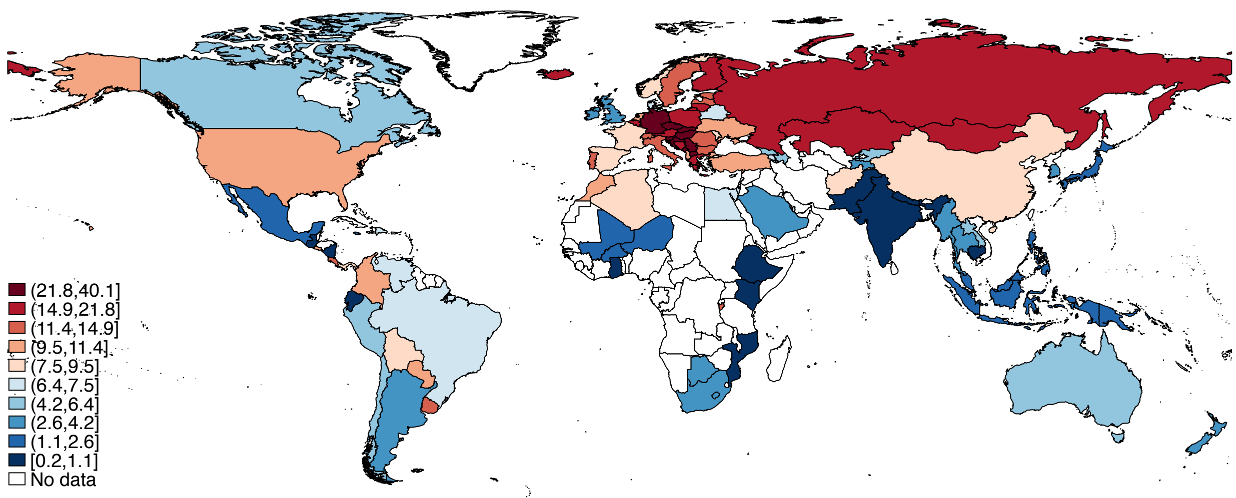
(OAS). The data provided by those four international organizations is long recognized for their transparency and expertise on various judicial indicators. Our contribution is to have created a unified database by: 1) merging the existing data and 2) verifying and supplementing the available data with the help of ministries of justice, supreme courts, national statistical offices and academic publications. We end up with an unbalanced panel of judges per capita for 107 countries from year 1970 to 2014. See section A.2 of the appendix for a complete description of the merging process, sources, definitions (table A3), and descriptive statistics (table A4).

Figure 2 gives an overview of our data by plotting the average number of judges per 100.000 inhabitants over the whole 1970-2014 period categorized by deciles. By this mean, we can see two things: first, the 107 countries for which we managed to collect at least one data point; second, the high cross-country variation of the number of judges per capita. The highest average density of judges for the 1970-2014 period is reached in Montenegro with 40.14 judges per 100.000 inhabitants; the lowest is found in Ethiopia with 0.24 judges per 100.000 inhabitants. Even within-continents the cross-country variation stays sizeable. For example, in Europe we find both countries belonging to the top decile (like Germany or Serbia) and countries belonging to the third lowest decile (like the United Kingdom or Ireland). The cross-country variations in terms of judges per capita are driven by differences in factors such as GDP per capita, political regime, legal origin, culture or ethnic composition of the population.[4] For instance, the legal origin is one variable that can help us understand the difference in the density of judges between two countries that have similar income per capita and similar political regime like Germany and the United Kingdom.

Figure 3 looks at the evolution of five-year averages of the density of judges across different regions. Even though during the last 45 years the world population has increased, the number of judges has increased even more. The world average of the density of judges has more than doubled, going up from 4.26 to 10.63 judges per 100.000 inhabitants between 1970 and 2014.

[4]See Appendix A.4. for more details on the determinants and correlates of the density of judges in our sample.

Figure 2: Average density of judges around the world between 1970-2014 (deciles)



Note: Country-level distribution of the number of judges per 100.000 inhabitants (averaged between 1970-2014). Each color represents a decile from the first (blue) to the tenth (red).

This doubling of the density of judges was achieved at a quite stable growth rate over the analyzed period.

Europe consistently displays the highest average density of judges among all continents, starting from 7.78 judges in 1970, to 21.63 judges per 100.000 inhabitants in 2010. Testing the difference of means between Europe and the rest of the world across the whole period we find that, on average, European countries have 10 more judges per 100.000 inhabitants than the rest of the world.

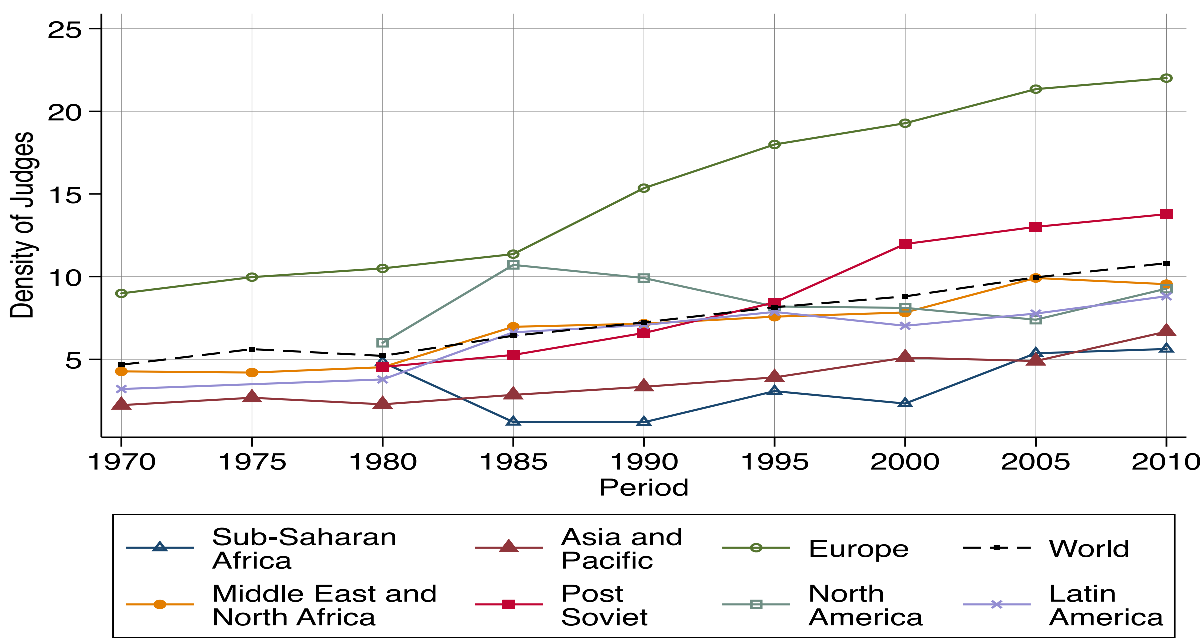
We document a remarkable increase in the number of judges per capita in the Post-Soviet countries since the 1990's.[5] Since 1980 the average number of judges in Post-Soviet countries has more than tripled going from 4.53 to 14.62 judges per 100.000 inhabitants in 2010. Post-Soviet countries have now the second highest density of judges in the world after Europe. Since the 1990's, an average Post-Soviet country has 2.9 more judges per 100.000 inhabitants than an average American country. One of the main explanations for the judicial staff increase in Post-Soviet countries after 1990, is the replacement of the soviet administration after the fall of USSR which had to meet legal needs following the transformation from

[5]In our sample, Post-Soviet countries are comprised of Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyz Republic, Latvia, Lithuania, Moldova, Russia and Ukraine.

planned to market economies (Dietrich; 2000; Murrell; 2001).

Finally, figure 3 highlights the sluggish evolution of the average density of judges in the other considered regions: Africa, Asia, the Middle-East, North America and Latin America. Even though North America (United States and Canada) always have a positive growth rate on the number of judges, the population growth rate is even higher in some periods (e.g. between 1985 and 1995), leading to an overall decrease in the average density of judges.

Figure 3: Average density of judges across regions between 1970-2014



Note: This figure plots trends in the density of judges (number of judges per 100.000 inhabitants) between 1970-2014. The graph covers 107 countries categorized into seven regions plus the World. The seven regions used are identical to the World Bank classification with the exception of the post-Soviet region: Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyz Republic, Latvia, Lithuania, Moldova, Russia, Tajikistan and Ukraine.

3 Empirical Strategy

Our goal is to empirically investigate the causal effect of ATJ (as proxied by the number of judges per capita) on economic development (as proxied by the growth rate of GDP per capita). In this section, we present our identification strategy and we discuss the diagnostic tests that we use throughout the paper to deal with endogeneity issues. We use five-year averages of all variables to smooth the short-run fluctuations and handle the annual gaps in the data. Moreover, changes in the judicial system are expected to have their full effect on the economy after a few years rather than immediately in one year of time.

To explain our estimation strategy, we take the following dynamic model as a starting point:

$$\ln\left(\frac{y_{i,t}}{y_{i,t-1}}\right) = \beta_1 \ln(y_{i,t-1}) + \beta_2 \ln(ATJ_{i,t-1}) + \alpha_i + \delta_t + \varepsilon_{i,t} \tag{1}$$

Where $y_{i,t}$ is GDP per capita in country i at time t, $ATJ_{i,t}$ is the number of judges per 100.000 inhabitants in country i at time t. α_i and δ_t are country and time fixed-effects, and $\varepsilon_{i,t}$ is the error term. Equation (1) corresponds to a beta-convergence model, which is a standard specification in the empirical growth literature since Islam (1995). If negative and above -1, the β_1 coefficient captures the speed at which an economy is converging to the common long-run GDP per capita level. By controlling for the dynamic component of GDP per capita, this model can isolate the effect of the other right hand-side variables on the steady-state of the economy. Our coefficient of interest is β_2 as it captures the short-run effect of ATJ on economic growth.

Estimating equation (1) by pooled OLS will generate inconsistent estimates. The key issue is the endogeneity of our variable of interest, ATJ. First, ATJ can affect economic growth (for example through contract enforcement) and be affected by economic growth (as citizens with higher income can more easily afford to go on court). This means that we are facing reverse causality. Second, access to justice and economic prosperity can be jointly affected by a third omitted variable (e.g. political regime), leading to a omitted variable bias. Thirdly, the dynamic panel specification leads to an asymptotic bias of order $1/T$ known as the Nickel bias (Nickell; 1981). In our case, this is a concern as we have a relatively short time dimension (in our estimations $T = 9$).

To address the aforementioned biases, we apply an instrumental variable approach. More specifically, we estimate equation (1) using a two-step difference Generalized Method of

Moments (GMM):

$$\Delta \ln\left(\frac{y_{i,t}}{y_{i,t-1}}\right) = \beta_1 \Delta \ln(y_{i,t-1}) + \beta_2 \Delta \ln(ATJ_{i,t-1}) + \Delta\delta_t + \Delta\varepsilon_{i,t} \tag{2}$$

In that form, equation (2) is estimated in the spirit of the seminal paper of Arellano and Bond (1991). By taking equation (1) in first difference and including time fixed-effects, we are able to control for all the time invariant country characteristics affecting both economic growth and ATJ such as: legal origin, culture, ethnic composition of the population, structural criminality, geography, etc. The inclusion of time fixed-effects allows to capture the world economy trends and the business cycle effects. The identification strategy relies on internal instrumentation: lagged levels of the right-hand side variables are used to instrument current differences. If the residuals are not serially correlated, difference GMM yields consistent estimates in a dynamic panel model with small T, helping to mitigate both the Nickel bias and the reverse causality issues. Treating ATJ as endogenous means that the first available instrument for $\Delta ATJ_{i,t-1}$ is $ATJ_{i,t-2}$ with the assumption that the level of ATJ observed at $t-2$ (i.e. 10 years before) is not correlated with current shocks of economic growth. We treat the log of GDP per capita as predetermined and time-period dummies as exogenous.[6] Specifically, we focus on the following moment conditions:

$$\mathbb{E}\left[(\varepsilon_{i,t} - \varepsilon_{i,t-1})\left(\ln(y_{i,t-j}), \ln(ATJ_{i,t-k})\right)\right] = 0 \qquad \text{for all } 1 \leq j \leq 4 \text{ and } k = 2 \tag{3}$$

The exact choice of instruments follows a trade-off between the moment condition used (which is most likely to hold using deeper lags) and the relevance condition (the fact that instruments should be sufficiently correlated with the instrumented variables), which is most likely to be satisfied with shorter lags.[7]

[6] In the growth empirics literature using GMM, there is no consensus on how to treat the initial level of GDP per capita: while some studies treat it as predetermined (DeJong and Ripoll; 2006); some treat it as endogenous (Hauk and Wacziarg; 2009; Voitchovsky; 2005). In the appendix table A7, we provide evidence that our results are qualitatively the same if we treat GDP per capita as endogenous.

[7] In appendix table A7 we provide evidence of the robustness of our benchmark specification to different moment conditions (i.e. different choice of internal instruments).

Difference GMM estimation of equation (2) constitutes our parsimonious benchmark specification and is used to explore further heterogeneous effects of ATJ on economic growth via the following specification:

$$\Delta \ln \left(\frac{y_{i,t}}{y_{i,t-1}} \right) = \beta_1 \Delta \ln(y_{i,t-1}) + \beta_2 \Delta \ln(ATJ_{i,t-1}) + \gamma' \Delta(\ln(ATJ_{i,t-1}) * \mathbf{\Theta}_{i,t-1}) + \Delta\delta_t + \Delta\varepsilon_{i,t} \tag{4}$$

Where $\mathbf{\Theta}_{i,t-1}$ is a vector of time-variant (e.g. level of democracy or education) or time-invariant variables (e.g. the legal origin or geographical area) interacted with ATJ in country i at time t. When estimating equation (4), we keep the same moment conditions for the interaction terms as those of ATJ.

To evaluate the quality of our internal instrumentation, we systematically report p-values of the Hansen (1982) J test and the Arellano-Bond test for AR2. These are the two standard GMM diagnostic tests for the quality of the instrumentations. The Hansen test is heteroskedasticity robust and it evaluates the joint validity of all the instruments. The Arellano-Bond test for AR2 evaluates the second order autocorrelation of residuals; this is required for a good instrumentation as the presence of second order autocorrelation would yield $ATJ_{i,t-2}$ to be an invalid instrument. Since Roodman (2009) and Windmeijer (2005), economists are aware of the problem of too many instruments when using GMM.[8] A common solution is to reduce the number of instruments below the number of individuals. In our estimations, we always keep the instrument count well below the number of countries in our sample. In particular, our benchmark specification uses a collapsed matrix of instruments: a method allowing to reduce significantly the instrument count without losing information.[9]

To evaluate further the quality of our internal instrumentation, we follow Bazzi and Clemens (2013) testing for underidentification and weakness of the GMM instrument matrix in a Two Stage Least Squares (2SLS) context. We first provide p-values for a test of

[8] In a difference GMM estimation without any restriction on the number of lags, the number of instruments is quadratic in T.

[9] In the collapsed form the matrix of instruments contains one column per lag, instead of one column per lag and time-period in the non-collapsed form.

underidentification based on the Lagrange Multiplier version of Kleibergen and Paap (2006) rk statistic (hereafter KP-LM). The KP-LM test evaluates if one or more of the canonical correlations is zero, providing a lower hurdle test for the weakness of instruments. Second, we provide p-values to test the weakness of instruments as implemented by Bazzi and Clemens (2013). In particular, this test uses a Wald F statistic based on the Kleibergen and Paap (2006) rk statistic and the critical values classified by Stock et al. (2005). The Kleibergen-Paap Wald test (hereafter KP-W) evaluates if the bias in the point estimate of our endogenous variable is greater than 30 percent of the OLS bias.

Another GMM estimator, the Blundell and Bond (1998) system GMM is often used in the economic literature. This estimator jointly estimates the difference equation with a level equation in which the levels of right hand side variables are instrumented with lagged differences. However, this estimator is not suitable in our case for two main reasons. Firstly, system GMM requires additional moment conditions to hold between the current level of the error term and the lagged differences. The crucial point here is that the error term in the level equation still contains the fixed-effect. Roodman (2009) shows that the additional moment condition is not satisfied if the individuals in the sample have different initial deviations from their long-run mean. As our sample is comprised of many heterogeneous economies (transition economies, newly industrialized, developed) this condition is not likely to hold. Secondly, the system GMM estimation can produce weak instruments too. Bazzi and Clemens (2013) show that in several published growth papers, the superiority of system GMM can be questioned and that the weakness of instruments should be directly tested: "In practice, most applications of system GMM simply assume that instruments are strong. We argue that instrument strength is an empirical question that can and should be directly tested". Other researchers advance a similar argument in the literature such as Hayakawa (2009) and Bun and Windmeijer (2010).[10]

[10]Bun and Windmeijer (2010) show that when the ratio of the variance of the country fixed-effect to the variance of the idiosyncratic shocks is higher than unity, the problem of weak instruments in the level equation of system GMM is amplified. Considering equation (2), we find a ratio of 2.6, indicating that system GMM is likely to produce weak instruments.

Finally, the potential non-stationarity of both ATJ (looking at figure 3) and of GDP per capita is not an identification issue in our setting for several reasons. First, contrary to panels with large T, unit-root is not a major concern in panels with relatively small T ($T = 9$ in our case). Second, we detrend our potentially non-stationary panel by estimating our benchmark specification in first differences using equation (2). Furthermore, the lagged dependent variable coefficient in equation (2) is significantly negative which is consistent with the stationarity assumption.

4 Results

The results are organized in two subsections. First, we present the main results obtained from the two-step difference GMM estimation and discuss their robustness. In the second subsection, we check for heterogeneous effects across geographical areas, income levels, legal origin, political regime and education.

4.1 Baseline Results and Robustness Checks

Table 1 presents our results based on equation (2) corresponding to a two-step difference GMM estimation on an unbalanced panel of 83 countries for the 1970-2014 period. Column 1 is our benchmark specification. Columns 2 to 4 show alternative difference GMM estimations. Columns 5 to 11 show the robustness of the benchmark specification to different subsamples. For each estimation, we report four diagnostic tests to keep track of the quality of our GMM estimations (AR2, Hansen, KP-LM and KP-W).

In our benchmark specification (column 1), we find that increasing ATJ by 1% causes an increase in the five-year growth rate of GDP per capita by 0.86 p.p. (0.17 p.p. annually), which is a sizeable effect. Our benchmark results are in line with other cross-country studies quantifying the effect of effective judicial institutions on economic growth. In particular, Voigt et al. (2015) in a cross-section of 100 countries find that increasing de facto judicial independence by one standard deviation leads to a 0.3 p.p. increase in annual GDP per capita growth. Melcarne and Ramello (2016) in a panel of 175 countries over 12 years using fixed-effect estimation, show that an extra year in judicial delays of private litigation lowers annual growth rate by over 1 p.p.. Our estimates are even more plausible as we tackle explicitly endogeneity issues.

In our benchmark specification, we deal with the issue of too many instruments inherent to GMM estimations by collapsing the instrument matrix: this allows us to keep the number of instruments well below the number of country in our sample as suggested by Roodman (2009). The internal instrumentation seems valid as both the AR2 and the Hansen tests are not rejecting their null hypothesis. In addition, following Bazzi and Clemens (2013) we give evidence that the set of instruments used does not suffer from underidentification as we can reject the null hypothesis of the KP-LM test at the 1% level, and the KP-W suggests strong enough instruments to withdraw a sizeable portion of the OLS bias as we can reject the null hypothesis at the 5% level.

Table 1: Main results and robustness checks

		Alternative GMM estimations			Robustness to subsamples				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Bench	NoColl	Ortho	OrthoNoColl	<2Mln Inh	Low Judge	High Judge	<4 Obs	<6 Obs
L.ln(GDPpc)	-0.705***	-0.692***	-0.765***	-0.659***	-0.699***	-0.684***	-0.713***	-0.688***	-0.636***
	(0.086)	(0.072)	(0.092)	(0.085)	(0.087)	(0.082)	(0.096)	(0.081)	(0.049)
L.ln(ATJ)	0.860***	0.458***	0.733***	0.370***	0.845***	0.865***	0.966***	0.804***	0.523***
	(0.216)	(0.141)	(0.189)	(0.099)	(0.218)	(0.204)	(0.255)	(0.202)	(0.104)
AR2	0.498	0.851	0.756	0.474	0.598	0.455	0.384	0.546	0.768
Hansen	0.344	0.125	0.197	0.122	0.434	0.453	0.376	0.660	0.908
KP-LM	0.001	0.041	0.019	0.178	0.002	0.000	0.005	0.001	0.011
KP F-stat	6.18	3.67	3.61	2.47	6.09	7.02	5.06	5.86	11.87
KP-W	0.022	0.225	0.224	0.850	0.024	0.009	0.064	0.030	0.000
Instruments	12	39	12	39	12	12	12	12	12
Countries (N)	83	83	83	83	71	75	75	49	22
Observations (NT)	241	241	273	273	209	231	223	202	121

Note: Table reports two-step difference GMM estimations with Windmeijer-corrected standard errors in parentheses. All regressions include period fixed effects. Our dependent variable is five-year average growth rate of GDP per capita. All the estimations keep the same collapsed instrument lag structure, treating the lag dependent as predetermined (instrumenting it from lag 1 to lag 4) and the variable of interest as endogenous (instrumenting it only with lag 2). The KP LM underidentification test uses a rank test procedure from Kleibergen and Paap (2006). KP F-stat is heteroskedasticity robust multivariate analogues to the first-stage F statistics. For the KP-W test, since critical values do not exist for the KP statistic, we follow the approach suggested by Baum et al. (2007) and use the Stock et al. (2005) 30 percent of the OLS bias critical values for the multivariate statistic. Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01.

In column 2 we do not collapse the matrix of instruments, which raises the instrument count from 12 in our benchmark to 39. The instrumentation appears to be of lower quality with respect to the benchmark: we are able to reject the lower hurdle of the KP-LM test only at the 5% level. The set of instruments used is clearly weaker as the KP-W test indicates that our estimates contain more than 30% of the OLS bias. In particular, the ATJ coefficient is almost two times smaller, suggesting a downward bias.[11]

Columns 3 and 4 show results using forward orthogonal deviation instead of first difference transformation. By subtracting the average of all future observations for each variable, this technique allows to get rid of the time-invariant country characteristics without losing too much observations in an unbalanced panel.[12] Column 3 shows the estimation using both forward orthogonal deviation and a collapsed matrix of instruments. The ATJ coefficient is in line with the magnitude and the significance of our benchmark specification. Contrary to the benchmark, the KP-W test indicates a problem of weak instruments. Using forward orthogonal deviation without collapsing the matrix of instruments (column 4), leads to a weaker instrumentation as we are not able to reject the null hypothesis of both the KP-LM and the KP-W tests. The ATJ coefficient is now two times smaller with respect to the benchmark, revealing a clear bias toward OLS.

To reassure on the fact that our results are not driven by a specific group of countries, we report in columns 5 to 11 the robustness of our benchmark specification across different subsamples. In column 5, we remove countries with less than two million inhabitants to test if our results are not driven by low populated countries. Indeed, for a given minimum size of the judicial system, low populated countries have a higher density of judges and consequently higher ATJ. After dropping the twelve least populated countries from our sample, the results remain very robust both in terms of magnitude, significance and quality of instrumentation.[13]

[11] Table A6 shows the OLS estimation of equation (2). In particular, column 2 depicts an ATJ coefficient almost 10 times smaller with respect to our benchmark.

[12] The first difference transformation can be demanding in terms of losing observations in an unbalanced panel. However, in our case the use of forward orthogonal deviation leads to a modest gain of 32 observations in columns 3 and 4 with respect to the benchmark.

[13] Countries bellow 2 million inhabitants in 1990 in our sample: Botswana, Cyprus, Estonia, Iceland, Luxembourg, Macedonia, Malta, Mauritius, Montenegro, Qatar, Slovenia, Trinidad and Tobago.

In columns 6 and 7, we check the validity of our results to the removal of top and the bottom decile in terms of the density of judges (8 countries). Removing the bottom decile (column 6) does not change the significance or the magnitude of the ATJ coefficient with respect to the benchmark.[14] The quality of the instrumentation is now even better as we are able to reject the null hypothesis of the KP-W test at the 1% level. Removing the top decile (column 7) increases the ATJ coefficient with respect to the benchmark while the quality of the instrumentation stays close to the benchmark.[15]

In columns 8 and 9, we remove countries with less than 4 and 6 points of data for the density of judges. By doing so, we move gradually to a more balanced panel and we can verify if our results are driven by the unbalanced nature of the data. In both estimations, the ATJ coefficient stays positive and highly significant. Even when we are dropping almost 75% of the countries and 50% of the observations in the more demanding estimation (column 9), the results stay in line with the benchmark specification both in terms of significance and the quality of the instrumentation.

In the appendix table A7, we run additional robustness checks by considering different moment conditions (i.e. different sets of internal instruments) compared to the benchmark specification reported in column 1. Overall, our results remain robust across different choices of instruments unless we use too deep lags (see discussion in section 3).

4.2 Heterogeneity of the Results

Table 2 reports the results of the two-step difference GMM estimation of equation (4) for the 1970-2014 period. Columns 1 through 6 present the heterogeneity of the impact across different regions, and columns 7 to 9 across different income levels.

In column 1 we find a significantly smaller effect of ATJ on economic growth in Europe compared to the rest of the world. The quality of the instrumentation is even higher than

[14] The eight dropped countries happen to be mostly low income countries: Cambodia, Ecuador, Ethiopia, Guatemala, India, Malaysia, Mexico and Nepal.

[15] The eight dropped countries are all Central European and Balkan countries: Croatia, Czech Republic, Hungary, Luxembourg, Macedonia, Montenegro, Serbia and Slovenia.

in the benchmark as we can reject the null hypothesis of weak instruments at the 1% level. Increasing ATJ by 1% rises the five-year GDP per capita growth rate by 0.45 p.p. on average in Europe, compared to 0.84 p.p. outside Europe. We find this result in line with the literature on court productivity determinants (Beenstock and Haitovsky; 2004; Dimitrova-Grajzl et al.; 2012, 2016; Grajzl and Silwal; 2017). As discussed in section 2.1, there are higher returns to GDP for countries with understaffed judiciaries or low ATJ. This is typically the case outside Europe since the other continents/regions have significantly lower density of judges (figure 3). This is in line with our general finding on diminishing marginal returns: the expected economic gain of increasing ATJ is higher for lower initial levels of ATJ.

When looking at other regions, we do not find significant heterogeneous effects of ATJ in North America, Latin America, Sub-Saharan Africa, Middle East and North Africa, and Asia and Pacific (column 2 through 6). This means that ATJ is equally important for development in those geographical areas compared to the rest of the world. In other words, continental heterogeneity does not seems to affect the impact of ATJ in this regions. This is in line with figure 3 showing similar levels of density of judges across the globe, except for Europe. The quality of the instrumentation stays stable relatively to the benchmark, except in column 2 and 6 where we are not able to reject the null of weak instruments.

Exploring the heterogeneity of the effect by income levels (columns 7 to 9), we do not find any significant difference in the effect of ATJ across poor or rich economies. On the other hand, the ATJ coefficient stays in line with the significance and the magnitude of the benchmark. The quality of the instrumentation is in line with the benchmark, except in column 9 where we cannot reject the null of weak instruments.

Table 3 presents the heterogeneity of the results by legal origins, political regime and human capital levels. Column 1 suggests that the positive impact of ATJ on economic growth is equally important in common law as in civil law countries. This is an interesting finding as we would have expected a significantly different effect due to their fundamental difference. Indeed, the comparative law literature describes common law as an adversarial

Table 2: Heterogeneity across regions and income levels

	Region						Income level		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	EU	N_AMER	LAC	SSA	MENA	APAC	low_midinc	up_mid_inc	high_inc
L.ln(GDPpc)	-0.738***	-0.699***	-0.702***	-0.706***	-0.711***	-0.767***	-0.618***	-0.709***	-0.695***
	(0.089)	(0.103)	(0.085)	(0.086)	(0.085)	(0.105)	(0.097)	(0.085)	(0.092)
L.ln(ATJ)	0.843***	0.856***	0.860***	0.860***	0.863***	0.778***	0.575***	0.790***	0.933***
	(0.199)	(0.224)	(0.212)	(0.216)	(0.221)	(0.220)	(0.209)	(0.235)	(0.233)
L.ln(ATJ) X Θ	-0.395**	-1.041	0.303	-0.509	0.121	0.296	0.817	0.069	-0.139
	(0.191)	(9.800)	(0.443)	(0.327)	(0.530)	(0.291)	(0.683)	(0.244)	(0.246)
AR2	0.883	0.508	0.447	0.496	0.486	0.929	0.398	0.545	0.381
Hansen	0.156	0.281	0.429	0.346	0.318	0.458	0.280	0.295	0.341
KP-LM	0.000	0.424	0.001	0.001	0.001	0.029	0.001	0.001	0.005
KP F-stat	6.69	0.89	4.87	4.98	4.65	2.40	5.99	5.04	3.86
KP-W	0.005	0.930	0.048	0.043	0.061	0.469	0.013	0.040	0.139
Instruments	13	13	13	13	13	13	13	13	13
Countries (N)	83	83	83	83	83	83	83	83	83
Observations (NT)	241	241	241	241	241	241	241	241	241

Note: Table reports two-step difference GMM estimations with Windmeijer-corrected standard errors in parentheses. All regressions include period fixed effects. Our dependent variable is five-year average growth rate of GDP per capita. Θ is our interaction term component which takes the value of the corresponding title of the column. All the estimations keep the same collapsed instrument lag structure by treating the lag dependent as predetermined (instrumenting it from lag 1 to lag 4) and the variable of interest together with its interaction term components are treated as endogenous (instrumenting it only with lag 2). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01.

dispute resolution system, where the judge has less investigative powers than lawyers, while civil law countries dispute resolution system is inquisitorial, giving more abilities to the judge to manage the procedure (Zweigert et al.; 1998). The ATJ coefficient stays in line with the benchmark both in terms of magnitude and significance. However, looking at the KP-W test we conclude at the presence of weak instruments. Therefore, the size of the effect remains uncertain.

In columns 2 to 4, we explore heterogeneity within the civil law legal family, which can be divided into three main categories: French, German and Scandinavian legal origin. As the corresponding interaction terms are statistically insignificant, we cannot conclude a differential effect of French or German legal origin compared to other legal families. We do find a significant difference for Scandinavian legal origin at 10% level but as the number of observations for this legal family is very limited in our sample, we cannot infer a general tendency for this result.[16] Moreover, the KP-W test indicates a bias in the coefficients

[16] Countries of Scandinavian legal origin in our sample are: Denmark, Finland, Iceland, Norway and Sweden. Data coverage for that group is limited as the average number of observations is 2. Moreover, half

towards OLS in columns 3 and 4.

In columns 5 and 6, we test for differences in the effect across the level of democracy and judicial corruption. In column 5, we use the Polity2 score for measuring democracy, which is interacted with ATJ and used as a separate control. Although the effect of ATJ stays close to the benchmark in terms of magnitude and significance, we do not find a significant interaction between Polity2 and ATJ. This result is surprising as we would have expected a higher effect of ATJ in democracies. Indeed, democracy is a sign of inclusive institutions and a higher trust in judiciary institutions. In column 6, we test if increasing ATJ in a corrupted judiciary leads a heterogeneous effect. Using a score on corrupted judicial decision, we do not find a significant heterogeneous effect of ATJ. The ATJ coefficient stays in line with the benchmark both in terms of significance and magnitude. Althought we can reject the null of underidentification, we cannot reject the null of the KP-W test indicating a bias towards OLS.

In the last three columns we check if there are differential effects of ATJ across different human capital levels, as proxied by the share of people having completed primary (column 7), secondary (column 8) and secondary or tertiary education (column 9). Along the three education dimensions we find similar results compared to the benchmark in both the magnitude and the significance of the ATJ coefficient, while we do not a find coefficient of the interactions with education levels. There is a strong concern for the quality of the instrumentation as we fail to lower hurdle of underidentification in columns 8 and 9 and we fail to reject the null of weak instruments in columns 7 through 9.

In the appendix table A8, we check if the effect of ATJ was significantly larger in some specific periods. Overall, we conclude that ATJ was equally important for GDP growth throughout the whole 45-year period between 1970 and 2014. In addition, we do not find significant differential effect comparing pre-1990 to post-1990 period, indicating that ATJ remains of equal importance for economic development in current times.

of the available observations are for the most recent period (2010-2014)

Table 3: Heterogeneity across legal origins, political regimes and education

	Common law	Civil law			Political regime		Education		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	legor_uk	legor_fr	legor_ge	legor_sc	polity2	ju_corrupt	lpc	lsc	lhc_lsc
L.ln(GDPpc)	-0.718***	-0.722***	-0.684***	-0.711***	-0.717***	-0.692***	-0.717***	-0.669***	-0.660***
	(0.080)	(0.075)	(0.086)	(0.091)	(0.091)	(0.106)	(0.078)	(0.094)	(0.112)
L.ln(ATJ)	0.786***	0.839***	1.008***	0.867***	0.828***	0.986***	0.817***	0.772***	0.836***
	(0.209)	(0.219)	(0.322)	(0.234)	(0.194)	(0.305)	(0.193)	(0.268)	(0.274)
L.ln(ATJ) X Θ	0.062	-0.058	-0.205	-1.965*	-0.006	-0.150	-0.003	0.002	0.002
	(0.451)	(0.220)	(0.244)	(1.037)	(0.008)	(0.406)	(0.005)	(0.004)	(0.004)
L.Θ					0.012	0.247	0.016	-0.019	-0.025
					(0.023)	(0.680)	(0.013)	(0.022)	(0.026)
AR2	0.565	0.573	0.383	0.522	0.709	0.837	0.747	0.483	0.564
Hansen	0.163	0.321	0.321	0.223	0.223	0.455	0.933	0.650	0.950
KP-LM	0.038	0.001	0.011	0.007	0.004	0.057	0.002	0.231	0.247
KP F-stat	2.10	5.57	3.69	3.31	3.36	1.30	3.32	0.89	0.83
KP-W	0.567	0.021	0.163	0.230	0.190	0.846	0.197	0.943	0.954
Instruments	13	13	13	13	14	14	14	14	14
Countries (N)	83	83	83	83	78	79	76	76	76
Observations (NT)	241	241	241	241	228	230	227	227	227

Note: Table reports two-step difference GMM estimations with Windmeijer-corrected standard errors in parentheses. All regressions include period fixed effects. Our dependent variable is five-year average growth rate of GDP per capita. Θ is our interaction term component which takes the value of the corresponding title of the column. All the estimations keep the same collapsed instrument lag structure by treating the lag dependent as predetermined (instrumenting it from lag 1 to lag 4) and the variable of interest together with its interaction term components are treated as endogenous (instrumenting it only with lag 2). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01.

5 Conclusion

The importance of ATJ today is revealed by the 2030 Agenda of Sustainable Development Goal of the United Nations (2015). Goal 16 states: "Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels". Despite this renewed interest, there is no systematic panel cross-country analysis trying to establish and quantify the economic impact of ATJ. This is important not only in achieving this goal but also other Sustainable Development Goals such as the goal 1: no poverty.

In this paper, we build a new database on the number of judges per capita in 107 countries from 1970 to 2014. This measure allows us to proxy ATJ across countries and through time, bridging the actual data limitation in the literature. Secondly, we tackle endogeneity issues using difference GMM estimation with internal instruments and we uncover diminishing marginal returns to GDP per capita growth by improving ATJ. This result is robust to various subsamples of countries. In a dynamic panel setting we find that increasing ATJ by 1% increases GDP per capita growth rate by 0.86 p.p., with dimishing marginal returns. Third, checking for heterogeneous effects across continents, we find that increasing ATJ in Europe (the region endowed with the highest density of judges worldwide), leads to an two times smaller effect which is in line with the diminishing marginal returns argument. In contrast to that finding, we find no effect heterogeneity by income level, legal origin, political regime, corruption of the judiciary or education levels.

Overall, our results suggest that ATJ is an important factor affecting economic prosperity. Identifying the exact channels through which ATJ affects economic growth is an important question for future research. Moreover, by further expanding the time-dimension we can learn more about the long-run effects of access to justice which is another interesting area for future research.

References

Acemoglu, D., Johnson, S. and Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth, *Handbook of economic growth* **1**: 385–472.

Aldashev, G., Chaara, I., Platteau, J.-P. and Wahhaj, Z. (2012). Using the law to change the custom, *Journal of Development Economics* **97**(2): 182–200.

Amirapu, A. (2017). Justice delayed is growth denied: The effect of slow courts on relationship-specific industries in india, *Technical report*, School of Economics Discussion Papers 1706.

Arellano, M. and 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.

Baum, C. F., Schaffer, M. E. and Stillman, S. (2007). Enhanced routines for instrumental variables/generalized method of moments estimation and testing, *The Stata Journal* **7**(4): 465–506.

Bazzi, S. and Clemens, M. A. (2013). Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth, *American Economic Journal: Macroeconomics* **5**(2): 152–86.

Beck, T., Levine, R. and Loayza, N. (2000). Finance and the sources of growth, *Journal of financial economics* **58**(1-2): 261–300.

Beenstock, M. and Haitovsky, Y. (2004). Does the appointment of judges increase the output of the judiciary?, *International Review of Law and Economics* **24**(3): 351–369.

Berkowitz, D., Pistor, K. and Richard, J.-F. (2003). Economic development, legality, and the transplant effect, *European Economic Review* **47**: 165 – 195.

Bielen, S., Marneffe, W., Grajzl, P. and Dimitrova-Grajzl, V. (2018). The Duration of Judicial Deliberation: Evidence from Belgium, *Journal of Institutional and Theoretical Economics (JITE)* **174**(2): 303–333.

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

Botero, J. C., Djankov, S., Porta, R. L., Lopez-de Silanes, F. and Shleifer, A. (2004). The regulation of labor, *The Quarterly Journal of Economics* **119**(4): 1339–1382.

Bun, M. J. and Windmeijer, F. (2010). The weak instrument problem of the system gmm estimator in dynamic panel data models, *The Econometrics Journal* **13**(1): 95–126.

Chemin, M. (2009a). Do judiciaries matter for development? evidence from india, *Journal of Comparative Economics* **37**(2): 230–250.

Chemin, M. (2009b). The impact of the judiciary on entrepreneurship: Evaluation of pakistan's access to justice programme, *Journal of Public Economics* **93**(1-2): 114–125.

Coviello, D., Ichino, A. and Persico, N. (2015). The inefficiency of worker time use, *Journal of the European Economic Association* **13**(5): 906–947.

DeJong, D. N. and Ripoll, M. (2006). Tariffs and growth: an empirical exploration of contingent relationships, *The Review of Economics and Statistics* **88**(4): 625–640.

Dietrich, M. K. (2000). *Legal and judicial reform in Central Europe and the former Soviet Union-voices from five countries*, The World Bank.

Dimitrova-Grajzl, V., Grajzl, P., Slavov, A. and Zajc, K. (2016). Courts in a transition economy: Case disposition and the quantity–quality tradeoff in bulgaria, *Economic Systems* **40**(1): 18–38.

Dimitrova-Grajzl, V., Grajzl, P., Sustersic, J. and Zajc, K. (2012). Court output, judicial staffing, and the demand for court services: Evidence from slovenian courts of first instance, *International Review of Law and Economics* **32**(1): 19–29.

Djankov, S., La Porta, R., Lopez-de Silanes, F. and Shleifer, A. (2002). The regulation of entry, *The quarterly Journal of economics* **117**(1): 1–37.

Djankov, S., McLiesh, C., Nenova, T. and Shleifer, A. (2003). Who owns the media?, *The Journal of Law and Economics* **46**(2): 341–382.

Djankov, S., McLiesh, C. and Shleifer, A. (2007). Private credit in 129 countries, *Journal of financial Economics* **84**(2): 299–329.

Feld, L. P. and Voigt, S. (2003). Economic growth and judicial independence: cross-country evidence using a new set of indicators, *European Journal of Political Economy* **19**(3): 497–527.

Galanter, M. (1981). Justice in many rooms: Courts, private ordering, and indigenous law, *The Journal of Legal Pluralism and Unofficial Law* **13**(19): 1–47.

Grajzl, P. and Silwal, S. (2017). The functioning of courts in a developing economy: evidence from nepal, *European Journal of Law and Economics* pp. 1–29.

Hammergren, L. (2014). Reforming the reformers: A different approach to access issues, *Institutional Competition between Common Law and Civil Law*, Springer, pp. 349–379.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica: Journal of the Econometric Society* pp. 1029–1054.

Hauk, W. R. and Wacziarg, R. (2009). A monte carlo study of growth regressions, *Journal of Economic Growth* **14**(2): 103–147.

Hayakawa, K. (2009). On the effect of mean-nonstationarity in dynamic panel data models, *Journal of Econometrics* **153**(2): 133–135.

Islam, N. (1995). Growth empirics: a panel data approach, *The Quarterly Journal of Economics* **110**(4): 1127–1170.

Jappelli, T., Pagano, M. and Bianco, M. (2005). Courts and banks: Effects of judicial enforcement on credit markets, *Journal of Money, Credit and Banking* pp. 223–244.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition, *Journal of econometrics* **133**(1): 97–126.

La Porta, R., Lopez-de Silanes, F. and Shleifer, A. (2008). The economic consequences of legal origins, *Journal of economic literature* **46**(2): 285–332.

La Porta, R., Lopez-de Silanes, F., Shleifer, A. and Vishny, R. (1999). The quality of government, *The Journal of Law, Economics, and Organization* **15**(1): 222–279.

La Porta, R., Lopez-de Silanes, F., Shleifer, A. and Vishny, R. W. (1997). Legal determinants of external finance, *Journal of finance* pp. 1131–1150.

La Porta, R., Lopez-de Silanes, F., Shleifer, A. and Vishny, R. W. (1998). Law and finance, *Journal of political economy* **106**(6): 1113–1155.

Levine, R., Loayza, N. and Beck, T. (2000). Financial intermediation and growth: Causality and causes, *Journal of monetary Economics* **46**(1): 31–77.

Mahoney, P. G. (2001). The common law and economic growth: Hayek might be right, *The Journal of Legal Studies* **30**(2): 503–525.

Melcarne, A. and Ramello, G. (2016). Justice delayed, growth denied: Evidence from a comparative perspective, *Unpublished manuscript* .

Murrell, P. (2001). Demand and supply in romanian commercial courts: Generating information for institutional reform, *Available at SSRN 280428* .

Nickell, S. (1981). Biases in dynamic models with fixed effects, *Econometrica: Journal of the Econometric Society* pp. 1417–1426.

Platteau, J.-P. (2015). *Institutions, social norms and economic development*, Routledge.

Pleasence, P., Balmer, N. J. and Sandefur, R. L. (2013). Paths to justice: a past, present and future roadmap.

Ramseyer, J. M. (2012). Talent matters: Judicial productivity and speed in Japan, *International Review of Law and Economics* **32**(1): 38–48.

Roodman, D. (2009). A note on the theme of too many instruments, *Oxford Bulletin of Economics and statistics* **71**(1): 135–158.

Rosales-López, V. (2008). Economics of court performance: an empirical analysis, *European Journal of Law and Economics* **25**(3): 231–251.

Stock, J. H., Yogo, M. and Andrews, D. W. (2005). Identification and inference for econometric models: Essays in honor of thomas rothenberg, *Cambridge Univ Press* .

United Nations (2015). Transforming our world: The 2030 agenda for sustainable development, *Resolution adopted by the General Assembly* .

Visaria, S. (2009). Legal reform and loan repayment: The microeconomic impact of debt recovery tribunals in india, *American Economic Journal: Applied Economics* **1**(3): 59–81.

Voigt, S. and El-Bialy, N. (2016). Identifying the determinants of aggregate judicial performance: taxpayers money well spent?, *European Journal of Law and Economics* **41**(2): 283–319.

Voigt, S., Gutmann, J. and Feld, L. P. (2015). Economic growth and judicial independence, a dozen years on: Cross-country evidence using an updated set of indicators, *European Journal of Political Economy* **38**: 197 – 211.

Voitchovsky, S. (2005). Does the profile of income inequality matter for economic growth?, *Journal of Economic growth* **10**(3): 273–296.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators, *Journal of econometrics* **126**(1): 25–51.

Woolf, H. (1996). *Access to Justice: final report to the Lord Chancellor on the Civil Justice System in England and Wales*, HM Stationery Office.

Zweigert, K., Kötz, H. and Weir, T. (1998). *Introduction to Comparative Law*, Clarendon Press.
URL: *https://books.google.be/books?id=OBuQQgAACAAJ*

Appendix

A.1. List of Variables

Table A1: Variable definitions and sources - common variables

Variable	Description and Definition	Source
PANEL A - Time variant		
ln(GDPpc)	GDP per capita in logs. Real GDP/capita at constant 2011 national prices (in mil. 2011US$).	Feenstra et al. (2015)
ln(ATJ)	Log of access to justice proxied by the density of judges. Density of judges is calculated as number of judges per 100.000 inhabitants, where the population data is from World Bank (2018).	Authors' calculation
polity2	The Polity2 score is based on the constraints placed on the chief executive, the competitiveness of political participation, and the openness and competitiveness of executive recruitment. The score ranges from -10 to +10, with higher values indicating stronger democratic institutions.	Marshall et al. (2002)
ju_corrupt	Judicial corruption decision measuring how often do individuals or businesses make undocumented extra payments or bribes in order to speed up or delay the process or to obtain a favorable judicial decision. Time variant dummy = 1 if value is higher than the median level in a given year.	Coppedge et al. (2018)
lpc	Percentage of population with primary complete education.	Barro and Lee (2013)
lsc	Percentage of population with secondary complete education.	Barro and Lee (2013)
lhc_lsc	Percentage of population with secondary and tertiary complete education.	Barro and Lee (2013)
PANEL B - Time invariant		
EU	Dummy = 1 for Europe countries.	World Bank
N_AMER	Dummy = 1 for North America countries.	World Bank
LAC	Dummy = 1 for Latin America countries.	World Bank
SSA	Dummy = 1 for Sub-Saharan Africa countries.	World Bank
MENA	Dummy = 1 for Middle East and North Africa countries.	World Bank
APAC	Dummy = 1 for Asia and Pacific countries.	World Bank
low_midinc	Dummy = 1 for Lower-middle-income economies.	World Bank
up_midinc	Dummy = 1 for Upper-middle-income economies.	World Bank
high_inc	Dummy = 1 for High-income economies.	World Bank
legor_uk	Dummy = 1 for Common law and = 0 Civil law otherwise.	La Porta et al. (2008)
legor_fr	Dummy = 1 for French civil law.	La Porta et al. (2008)
legor_ge	Dummy = 1 for German civil law.	La Porta et al. (2008)
legor_sc	Dummy = 1 for Scandinavian civil law.	La Porta et al. (2008)
access_and_afford_just	People can Access and Afford Civil Justice. Measures the accessibility and affordability of civil courts, including whether people are aware of available remedies; can access and afford legal advice and representation; and can access the court system without incurring unreasonable fees, encountering unreasonable procedural hurdles, or experiencing physical or linguistic barriers.	The World Justice Project (2019)

Table A2: Summary Statistics

Variable	Mean	S.D.	Min.	Max.	Obs.
ln(GDPpc)	9.60	0.85	6.36	11.82	241
ln(ATJ)	2.07	0.92	-1.26	3.94	241
polity2	6.94	5.05	-10.00	10.00	230
ju_corrupt	0.32	0.47	0.00	1.00	234
lpc	19.34	12.06	0.04	51.01	227
lsc	29.47	14.84	2.92	71.91	227
lhc_lsc	41.50	18.15	3.54	84.69	227
EU	0.39	0.49	0.00	1.00	241
N_AMER	0.02	0.16	0.00	1.00	241
LAC	0.13	0.34	0.00	1.00	241
SSA	0.02	0.13	0.00	1.00	241
MENA	0.08	0.28	0.00	1.00	241
APAC	0.24	0.43	0.00	1.00	241
low_midinc	0.10	0.29	0.00	1.00	241
up_midinc	0.28	0.45	0.00	1.00	241
high_inc	0.61	0.49	0.00	1.00	241
legor_uk	0.21	0.41	0.00	1.00	241
legor_fr	0.46	0.50	0.00	1.00	241
legor_ge	0.28	0.45	0.00	1.00	241
legor_sc	0.04	0.20	0.00	1.00	241

A.2. A New Database on Judges: Method and Sources

We construct a new database on number of judges by collecting and pooling different official sources: international organizations, ministries, statistical offices and academic publications. The coverage of each database can be complementary or overlapping as they focus on different countries and periods.[17] Therefore, our contribution is to have collected and unified different data sources in one database.

During the merging process, we have followed 3 main criteria:

1. Maximize the coverage for a given country with one dataset. This criterion enables us to minimize the measurement error inherent to different definitions or counting

[17] To illustrate the complementary case between multiple datasets, lets consider the example of France: CEPEJ provides data for France and for European countries from 2003 to 2014. The ICPSR provides data for years 1982 and 1984; whereas Pistor et al. (1999) delivers the first available data point for year 1977. As a result, we use all three sources as they are complementary to each other. To illustrate the overlapping case between multiple datasets, lets consider the example of Belgium: CEPEJ provides data for Belgium and for European countries from 2003 to 2014. However, the Ministry of Justice of Belgium provided us with their official statistics from 1970 to 2014. As a result of overlapping between the two datasets, we decided to use only the data from the Ministry of Justice.

methods while merging different sources. This rule is important as it minimizes fake evolutions in a country series.

2. Control for unrealistic change between the two closest merged country-year data points. If we observe jumps in the number of judges (annual change larger than 50%) between the two datasets, we do not merge these datasets. There could be a judicial reform at play that radically changes the number of judges for the exact country-year observation when we merge different datasets, however this can only be the case for very few countries. There is a significantly higher probability of large changes in values when merging different dataset than it is to observe in the same period a drastic judicial reform.

3. Rank the importance of datasets when merging. In the case of overlapping data with similar country-year coverage, we chose the more preferred database according to the specified list: Ministries of Justice and Supreme Courts, National Statistical Offices, international organizations and academic publications. We prefer national sources given their responsibility to issue first such statistics. Moreover, we favor more recent datasets as they are more likely to be precise since they can contain corrected estimates on the number of judges for previous years.

Overall, our measure of the number of judges is available for 107 countries and covers the period 1970-2014. This is an unbalanced panel, meaning that we cannot follow each country-year observation in our cover period. To obtain the density of judges, we divide the number of judges by population as reported by World Bank (2018). We then calculate five-year averages to smooth the short-run fluctuations and handle the annual gaps in the data. Table A3 details the sources, definitions, countries and periods cover by each used source of data. Table A4 provides descriptive statistics.

Table A3: Number of Judges: Sources, Definitions and Coverage

Source	Definition	Sample used
PANEL A - International Organizations		
ICPSR (2010) (United Nations Surveys of Crime Trends and Operations of Criminal Justice Systems)	Professional judges or magistrates may be understood to mean both full-time and part-time officials authorized to hear civil, criminal and other cases, including in appeal courts, and make dispositions in a court of law.	67 countries, 1973-2014 (Online database and rapports)
CEPEJ (2016) (Council of Europe European Commission for the efficiency of justice)	Total number of professional judges in all types of courts. Professional judges are those who have been trained and who are paid as such and whose main function is to work as a judge and not as a prosecutor; the fact of working full-time or part-time has no consequence for their status. It does not include the court clerks that exist in some member states.	34 European countries 2002-2014 (2002-2009 rapports and 2010-2014 online database)
UNODC (2016) (United Nations Office on Drugs and Crime)	Professional Judges or Magistrates means both full-time and part-time officials authorized to hear civil, criminal and other cases, including in appeal courts, and to make dispositions in a court of law. Includes also authorized associate judges and magistrates.	24 countries, 2003-2015 (Online database and rapports)
OAS (2016) (Organization of American States)	Professional judges or magistrates understood as both full-time and part-time officials authorized to hear civil, criminal and other cases, including in appeal courts, and to make dispositions in a court of law. Including authorized associate judges and magistrates.	17 countries, 2003-2014 (Online database and rapports)
EUROSTAT (2016) (European Statistical Office)	Both full-time and part-time officials authorized to hear criminal and civil cases, including in appeal courts, and to make dispositions in a court of law. Authorized associate judges and magistrates are included.	7 European countries, 2002-2013 (Online database and rapports)
PANEL B - Public Institutions		
Ministries of Justice and Supreme Courts	Data for: Botswana, Belgium, Croatia, Estonia, Finland, Ireland, Latvia, Mali, Malta, New Zealand, Niger, Poland, Romania, Slovenia, South Korea, Sweden, Uruguay, Venezuela	18 countries, 1970-2014
National Statistical Offices	Data for: Australia, Armenia, Qatar and United States	4 countries, 1980-2014
PANEL C - Publications		
Albers (2003)	Data for: Armenia, Botswana, Cambodia, Ghana, Mexico, Mozambique, Pakistan, Papua New Guinea, Peru, Trinidad and Tobago	10 countries, 2000
Pistor et al. (1999)	Data for: China, France, Germany, Japan, India, South Korea and Malaysia	7 countries, 1970-1995
Calleros-Alarcón (2008)	Data for: Bolivia, Brazil, Paraguay and Venezuela	4 countries, 1993-2001
Schmiegelow and Schmiegelow (2014)	Data for: Cambodia, Laos and Tajikistan	3 countries, 1995-2011
Contini (2000)	Data for: Austria and Spain	2 countries, 1980-1990
Dakolias (1999)	Data for Ecuador	1 country, 1998
Kühn (2011)	Data for Poland	1 country, 1981
Turner (2009)	Data for Serbia	1 country, 2002
IMF (2012)	Data for Burundi	1 country, 2005-2010

Note: Overlapping coverage implies prioritization of data sources by its level of reliability and allows for cross-validation of data.

Table A4: Number of judges per 100.000 inhabitants - 107 countries, 1970-2014

Country	Mean	S.D.	Min.	Max.	Country	Mean	S.D.	Min.	Max.
Afghanistan	8.95	.	8.95	8.95	Kyrgyz Republic	6.18	0.64	5.32	6.73
Albania	11.42	2.12	8.53	13.27	Laos	5.92	.	5.92	5.92
Algeria	8.72	3.53	3.89	12.32	Latvia	14.91	8.42	6.35	26.05
Argentina	3.6	1.87	1.44	4.75	Lithuania	15.6	8.42	5.39	25.49
Armenia	6.7	0.8	5.95	7.54	Luxembourg	34.43	5.94	25.85	39.27
Australia	4.82	0.18	4.64	5	Macedonia	28.63	4.57	20.79	31.52
Austria	19.71	0.95	18.38	21.04	Malaysia	1.31	0.37	0.83	1.67
Azerbaijan	4.66	1.78	2.52	6.46	Mali	2.58	.	2.58	2.58
Bahamas	9.16	1.49	8.11	10.21	Malta	6.08	2.71	2.85	10.09
Bahrain	11.44	2.27	9.83	13.05	Mauritius	4.66	1.15	3.45	5.73
Barbados	8.2	0.42	7.9	8.5	Mexico	2.15	1.88	0.83	4.31
Belarus	7.43	2.68	4.64	10.27	Moldova	8.9	3.07	4.69	12.45
Belgium	19.72	2.68	16.28	22.9	Mongolia	15.77	0.7	14.96	16.24
Bolivia	8.85	1.35	7.4	10.05	Montenegro	40.14	1.15	39.46	41.47
Bosnia and Herzegovina	21.75	4.25	18.25	26.47	Morocco	10.22	.	10.22	10.22
Botswana	2.82	1.22	2.1	4.22	Mozambique	0.97	.	0.97	0.97
Brazil	7.46	0.99	6.13	8.35	Myanmar	2.72	0.23	2.46	2.86
Bulgaria	14.22	10.48	2.82	30.43	Nepal	1	0.19	0.81	1.2
Burkina Faso	1.81	0.12	1.73	1.9	Netherlands	9.98	3.76	5.77	14.48
Burundi	14.41	6.77	6.62	18.88	New Zealand	3.77	0.87	2.18	4.86
Cambodia	1.11	0.71	0.47	1.86	Nicaragua	0.94	0.34	0.7	1.19
Canada	5.29	3.1	0.68	7.4	Niger	1.84	0.14	1.75	1.94
Chile	5.12	3.19	2.51	10.54	Norway	9.07	1.96	6.67	11.12
China	8.75	6.05	2.58	16.62	Pakistan	1.05	.	1.05	1.05
Colombia	9.93	0.37	9.4	10.34	Panama	7.51	0.83	6.04	8.42
Costa Rica	14.28	7.04	4.98	25.48	Papua New Guinea	2.48	.	2.48	2.48
Croatia	36.7	9.43	22.24	44.55	Paraguay	10.39	0.16	10.27	10.51
Cyprus	6.9	1.93	4.73	9.31	Peru	4.73	2.13	2.99	7.58
Czech Republic	28.55	0.97	27.43	29.19	Philippines	2.17	0.47	1.73	2.85
Denmark	6.68	0.19	6.47	6.83	Poland	17.29	6.7	8.62	26.04
Dominican Republic	6.92	0.09	6.86	6.99	Portugal	12.13	4.93	5.34	18.8
Ecuador	0.97	0	0.97	0.97	Qatar	7.39	1.9	5.37	9.62
Egypt	7.11	3.7	4.5	9.73	Romania	13.45	6.1	5.22	20.56
El Salvador	10.75	0.08	10.69	10.8	Russia	15.98	8.41	6.16	22.91
Estonia	13.8	4.62	5.13	16.86	Saudi Arabia	3.17	0.17	3.05	3.3
Ethiopia	0.25	0.08	0.16	0.3	Serbia	34.54	2.56	32.32	37.35
Finland	16.44	1.51	13.73	18.07	Singapore	1.82	0.56	1.08	2.5
France	9.51	1.68	6.56	10.62	Slovak Republic	22.59	2.6	18.76	25.34
Georgia	6.44	0.99	4.76	7.19	Slovenia	38.97	10.83	25.82	51.51
Germany	22.04	3.13	17.89	26.08	South Africa	3.74	0.55	3.03	4.22
Ghana	0.93	.	0.93	0.93	South Korea	2.76	1.45	1.28	5.37
Greece	17.51	7.77	9.7	31.23	Spain	8.6	2.04	5.07	10.85
Guatemala	0.67	0.02	0.66	0.69	Sweden	11.49	0.73	10.78	12.37
Hungary	22.34	5.97	12.66	28.55	Switzerland	13.15	2.58	10.57	15.41
Iceland	16.23	1.27	15.14	17.79	Tajikistan	3.19	2.36	0.47	4.63
India	1.06	0.08	1	1.15	Thailand	4.23	1.71	2.24	6.35
Indonesia	1.63	0.03	1.6	1.65	Trinidad and Tobago	6.68	0.94	5.6	7.34
Ireland	2.61	0.49	2.07	3.22	Turkey	9.89	0.7	9.19	11.02
Israel	6.81	1	5.55	8.1	Ukraine	9.68	5.35	4.37	17.24
Italy	12	1.33	10.74	14.67	United Kingdom	2.74	0.94	1.57	3.57
Jamaica	2.57	1.01	1.07	3.25	United States	10.12	0.75	9.27	11.32
Japan	2.44	0.2	2.28	2.88	Uruguay	13.19	0.97	11.45	14.4
Kazakhstan	14.99	1.41	13.38	15.97	Venezuela	6.5	0.36	6.1	6.8
Kenya	1	0.24	0.83	1.17					

Note: List of all 107 countries with at least one 5-year period observation on the number of judges per 100.000 inhabitants. Countries with standard deviation = (.) have one 5-year period observation.

A.3. Density of Judges and Alternative Measures of ATJ

In this section, we present additional evidence for the validity of our macro-level proxy of ATJ: the number of judges per capita. Figure A1 shows the partial effects of the log density of judges across four different measures of ATJ, conditionnal on log GDP per capita. The log density of judges is averaged across the whole 2000-2014 period to maximize sample size. All the conditional correlation results presented in figure A1 are robust to additional controls such as the democracy/autocracy indicators from Polity2.[18]

Panel a. shows that the positive partial correlation between ATJ and the log density of judges shown in figure 1 is robust to using alternative years. In particular, we test the other extreme by taking the earliest available data in *The World Justice Project*, 2012 instead of 2019.

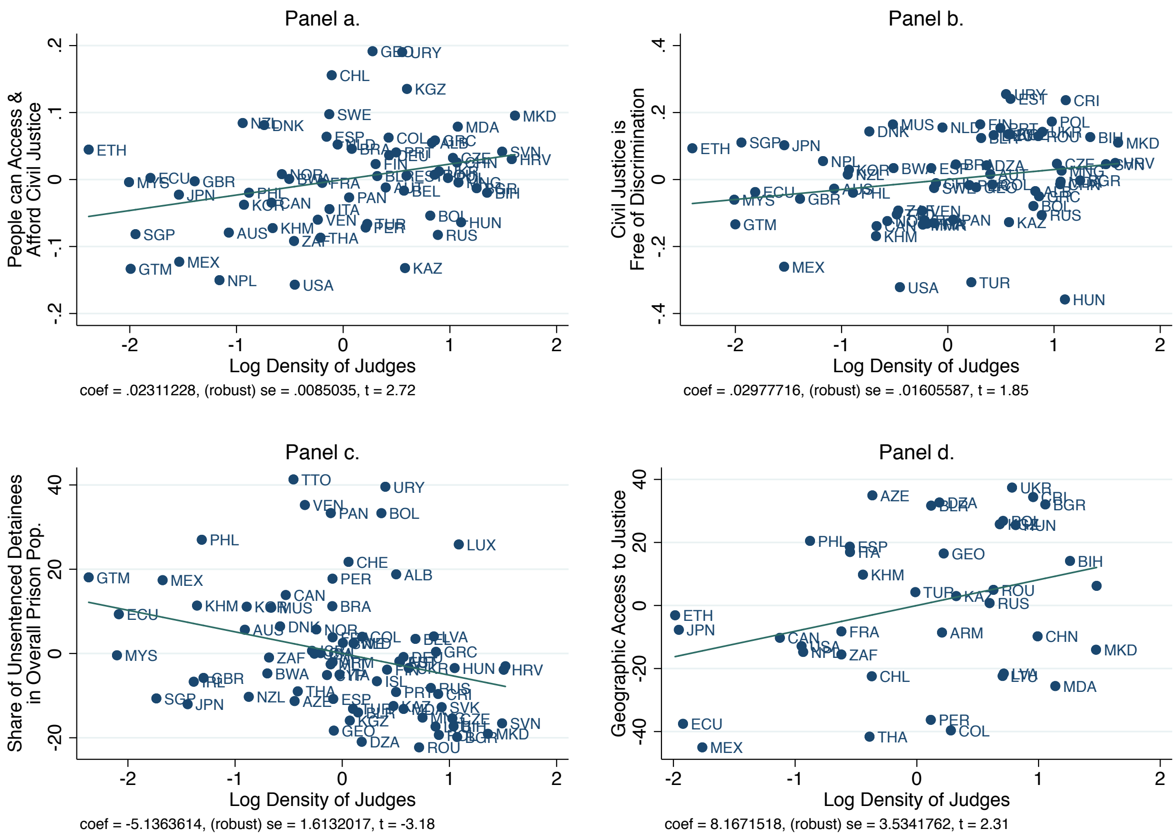
Panel b. tests the link between the log density of judges and the presence of discrimination along the judicial process (dimension 2 of ATJ, see section 2.1.) as measured by the 2019 edition of *The World Justice Project*. In particular, we use the sub-factor 7.2 measuring "whether the civil justice system discriminates in practice based on socio-economic status, gender, ethnicity, religion, national origin, sexual orientation, or gender identity." We find a positive and significant correlation confirming that our proxy is able to capture dimension 2 of ATJ. It should be noted that the relationship is hampered by Hungary which is clearly an outlying country in that case.

Panel c. tests the link between our proxy and the overall efficiency of the judicial system as measured by the share of unsentenced detainees in overall prison population. This is the official indicator chosen by the United Nations to track progress on the Rule of Law side of the Sustainable Development Goal 16.3. We find a robust negative correlation, meaning that the density of judges is positively associated with overall judicial efficiency.

Panel d. shows the partial correlation between our proxy and geographical ATJ (dimension 3 of ATJ, see section 2.1). In particular, we use a specific variable in the Global Integrity Index of 2007 and 2008 constructed by the non-profit organization Global Integrity asking whether "in practice, all citizens have access to a court of law, regardless of geographic location." Even with that noisy measure and a sample reduce to 38 countries, we are able to find a significant positive relationship between that measure of geographical accessibility to justice and the density of judges.

[18]Results available upon request

Figure A1: The Partial Effect of the Density of Judges on Alternative Access to Justice Measures



Note: This figure plots the relationship between the log density of judges (averaged between 2000-2014) and four measures of the accessibility and efficiency of the judicial system: the access to civil justice score of *The World Justice Project* in year 2012 (panel a.), the civil justice free of discrimination score of *The World Justice Project* in year 2019 (panel b.), the share of unsentenced detainees in overall prison population from UNODC in year 2016 (panel c.) and a score describing geographical access to justice from *Global Integrity* in year 2007-2008. This adjusted partial residual plots are based on OLS regressions where the log GDP per capita is used as a control variable and with robust standard errors.

A.4. Determinants and correlates of the Density of Judges

In this section, we look at the cross-country determinants and correlates of the density of judges. We estimate the following equation by OLS in an unbalanced panel setting:

$$\ln(Judgespc)_{i,t} = \alpha + \beta_1 \ln(y_{i,t}) + \beta_2 LO_i + \gamma' \boldsymbol{X}_{i,t} + \delta_t + \varepsilon_{i,t} \tag{5}$$

Where the dependent variable $Judgespc$ is the number of judges per 100.000 inhabitants in country i at time t, $y_{i,t}$ is GDP per capita in country i at time t, LO is the legal origin of country i, $\boldsymbol{X}_{it}$ is a vector of other explanatory variables in country i at time t, δ_t is a time fixed-effect and $\varepsilon_{i,t}$ is an error term. Each period t corresponds to a five-year average. Depending on the specification, $\boldsymbol{X}_{i,t}$ will contain the following variables: Polity2 score of democracy, the share of population having completed tertiary education, absolute latitude, percentage of population at risk of contracting malaria, ethnic fractionalization, percentage of population of European descent and share of population of a given religion. The mentioned variables are good candidates for explaining the cross-country variance of the density of judges and are widely used in the macroeconomic literature. Their sources and definitions are available in table A2.

The main finding of table A5 is that the legal origin is a robust predictor of the density of judges. We find that civil law countries of either French, German or Scandinavian legal origin, have significantly higher density of judges than comparable common law countries. Based on our preferred specification (column 3), being a French legal origin country increases the density of judges by 62% on average and being a German legal origin country more than double the density of judges, ceteris paribus. Moreover, inside the civil law legal family, we find that German legal origin countries have significantly higher density of judges than comparable French legal origin countries. These findings are consistent with the comparative law literature describing common law as an adversarial dispute resolution system, where the judge has less investigative powers than lawyers, while civil law countries dispute resolution system is inquisitorial, giving more power to the judge (Zweigert et al.; 1998). Common law legal system requires therefore less judges and more lawyers than civil law.[19] Differences inside the civil law family can be explained by specificities of the German legal origin in terms of judicial organisation (Schmiegelow and Schmiegelow; 2014). Finally, we confirm figure 3 showing that the share of population of European descent is a strong predictor of the density of judges.

Column 1 shows a positive and significant relationship between each of the three civil

[19]Explaining the cross-country differences in the number of lawyers, Massenot (2012) finds that common law and French civil law countries have more lawyers than German and Scandinavian legal origins countries.

law legal origin dummies (French, German and Scandinavian) and the log of the density of judges, controlling for the log of GDP per capita and year fixed-effects. Adding a measure of democracy and tertiary completed education (column 2) or a full set of macro controls proxying temperature, disease environment, colonisation patterns, ethnic composition of the population or cultural traits (column 3) does not alter the significance of the relationship found, except for Scandinavian legal origin coefficient which is still significant at the 10% level. Concerned by the possibility that post-soviet countries have changed their legal origin after the fall of the Berlin Wall (La Porta et al.; 2008), columns 4 to 6 replicate the first three columns taking the 1990-2014 period. Consistent with our previous findings, we confirm the highly significant and sizeable effect found for the three civil law legal origins. Performing a Wald test in each of the 6 mentioned columns, we are able to reject the hypothesis that the three civil law legal origin dummies are jointly insignificant at the 1% level, confirming the significance of the effect found. In columns 3 and 6, testing the equality of coefficients between the three civil legal origin families, we can reject the null hypothesis only when comparing German and French legal origins. It means that, among the civil law legal family, German legal origin countries enjoy on average more judges than their French counterpart, while Scandinavian legal origin countries are the in between category.

Among the control variables, one other important factor explaining the cross-country variance in the density of judges is the share of population of European descent (columns 3 and 6). This reflects our previous finding figure 3 as Europe is the host of nations with the highest density of judges worldwide. Furthermore, outside Europe this variable reflects the link between the colonization pattern and current institutions (Acemoglu et al.; 2005).

Table A5: Determinants and Correlates of the Density of Judges

	(1) 1970-2014	(2) 1970-2014	(3) 1970-2014	(4) 1990-2014	(5) 1990-2014	(6) 1990-2014
ln(GDPpc)	0.455***	0.346***	0.122	0.455***	0.364***	0.123
	(0.095)	(0.085)	(0.100)	(0.096)	(0.089)	(0.119)
French Legal Origin	0.919***	0.950***	0.623***	0.990***	0.995***	0.604***
	(0.162)	(0.154)	(0.211)	(0.160)	(0.154)	(0.217)
German Legal Origin	1.243***	1.276***	1.037***	1.425***	1.464***	1.117***
	(0.256)	(0.289)	(0.260)	(0.229)	(0.261)	(0.264)
Scandinavian Legal Origin	0.742***	0.729***	0.728*	0.734***	0.686***	0.678*
	(0.264)	(0.274)	(0.370)	(0.243)	(0.253)	(0.346)
Polity2		0.024**	0.011		0.018	-0.006
		(0.012)	(0.015)		(0.014)	(0.019)
% of Tertiary Completed		0.007	0.001		0.008	0.000
		(0.012)	(0.011)		(0.012)	(0.010)
Absolute latitude			-0.000			0.000
			(0.007)			(0.007)
% of pop. at risk of contracting malaria			-0.349			-0.490
			(0.440)			(0.409)
Ethnic fractionalization			0.336			0.253
			(0.350)			(0.369)
% of population of European descent			1.155***			1.180***
			(0.326)			(0.324)
Share of Protestants in the population			-0.004			-0.005
			(0.004)			(0.004)
Share of Roman Catholics in the population			-0.001			-0.001
			(0.002)			(0.003)
Share of Muslims in the population			0.004			0.003
			(0.003)			(0.003)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
adj. R-sq	0.43	0.45	0.60	0.41	0.43	0.60
Countries (N)	104	92	89	104	92	89
Observations (NT)	465	416	407	365	323	314

Note: the dependent variable is the number of judges per 100k inh. Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01.

A.5. Supplementary Results

Table A6: Pooled OLS and Fixed-effects Results

	OLS (1) 1970-2014	FE (2) 1970-2014	OLS (3) 1990-2014	FE (4) 1990-2014
L.ln(GDPpc)	-0.045***	-0.374***	-0.054***	-0.623***
	(0.009)	(0.071)	(0.012)	(0.106)
L.ln(ATJ)	0.005	0.089*	0.012	0.180***
	(0.010)	(0.045)	(0.008)	(0.051)
_cons	0.585***	3.382***	0.543***	5.572***
	(0.084)	(0.626)	(0.113)	(0.969)
Time FE	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	Yes
Adj.-R2	0.18	0.36	0.23	0.51
Countries (N)	83	83	83	83
Observations (NT)	356	356	260	260

Note: Our dependent variable is five-year average growth rate of GDP per capita. Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01.

Table A7: Robustness of benchmark to different moment conditions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
L.ln(GDPpc)	-0.705***	-0.709***	-0.708***	-0.703***	-0.702***	-0.560***	-0.702***
	(0.086)	(0.086)	(0.085)	(0.087)	(0.086)	(0.151)	(0.086)
L.ln(ATJ)	0.860***	0.849***	0.805***	0.877***	0.872***	0.907***	0.832***
	(0.216)	(0.217)	(0.217)	(0.217)	(0.214)	(0.253)	(0.192)
First_ln(GDPpc)	1	1	1	1	1	6	1
Last_ln(GDPpc)	4	4	3	4	4	7	5
First_ln(ATJ)	2	2	2	3	4	2	6
Last_ln(ATJ)	2	3	3	4	5	3	7
AR2	0.498	0.515	0.553	0.493	0.490	0.533	0.524
Hansen	0.344	0.435	0.391	0.411	0.504	0.820	0.532
KP-LM	0.001	0.002	0.006	0.002	0.002	0.001	0.003
KP F-stat	6.18	5.09	4.80	4.83	4.80	4.68	4.04
KP-W	0.022	0.055	0.082	0.072	0.074	0.091	0.153
Instruments	12	13	12	13	13	11	14
Countries (N)	83	83	83	83	83	83	83
Observations (NT)	241	241	241	241	241	241	241

Note: Table reports two-step difference GMM estimations with Windmeijer-corrected standard errors in parentheses. All regressions include period fixed effects. Our dependent variable is five-year average growth rate of GDP per capita. All the estimations report different instrument structure for each of the two right hand side variables. The matrix of instruments is collapsed across all estimations. Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01.

Table A8: Heterogeneity by Different Time-periods

	(1) period1	(2) period2	(3) period3	(4) period4	(5) period5	(6) period6	(7) period7	(8) period8	(9) post1990
L.ln(GDPpc)	-0.707***	-0.712***	-0.708***	-0.708***	-0.700***	-0.707***	-0.705***	-0.705***	-0.712***
	(0.087)	(0.087)	(0.089)	(0.085)	(0.088)	(0.081)	(0.084)	(0.087)	(0.083)
L.ln(ATJ)	0.857***	0.846***	0.867***	0.865***	0.889***	0.901***	0.843***	0.859***	0.872***
	(0.217)	(0.217)	(0.222)	(0.213)	(0.236)	(0.216)	(0.212)	(0.218)	(0.210)
L.ln(ATJ) X Θ	0.020	0.074	-0.078	0.026	0.027	-0.043	0.016	-0.002	-0.032
	(0.045)	(0.048)	(0.064)	(0.027)	(0.032)	(0.028)	(0.026)	(0.051)	(0.049)
AR2	0.494	0.464	0.412	0.522	0.563	0.651	0.525	0.507	0.572
Hansen	0.348	0.309	0.315	0.373	0.364	0.498	0.367	0.194	0.365
KP-LM	0.001	0.001	0.001	0.001	0.006	0.001	0.002	0.000	0.001
KP F-stat	6.12	4.95	4.76	5.28	4.30	4.90	4.61	6.11	4.88
KP-W	0.014	0.044	0.054	0.030	0.089	0.046	0.064	0.014	0.048
Instruments	13	13	13	13	13	13	13	12	13
Countries (N)	83	83	83	83	83	83	83	83	83
Observations (NT)	241	241	241	241	241	241	241	241	241

Note: Table reports two-step difference GMM estimations with Windmeijer-corrected standard errors in parentheses. All regressions include period fixed effects. Our dependent variable is five-year average growth rate of GDP per capita. Θ is our interaction term component which takes the value of the corresponding title of the column. All the estimations keep the same collapsed instrument lag structure by treating the lag dependent as predetermined (instrumenting it from lag 1 to lag 4) and the variable of interest together with its interaction term components are treated as endogenous (instrumenting it only with lag 2). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01.

Appendix references (not cited in the main text)

Albers, P. (2003). Evaluating judicial systems a balance between variety and generalisation. *European commission for the efficiency of justice (CEPEJ)*.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104:184–198.

Calleros-Alarcón, J. C. (2008). *The unfinished transition to democracy in Latin America*. Routledge.

CEPEJ (2016). Council of europe european commission for the efficiency of justice. Retrieved from `https://www.coe.int/en/web/cepej`.

Contini, F. (2000). European database on judicial systems. *European Research Network on Judicial Systems*.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., et al. (2018). Varieties of democracy (v-dem) database v8.

Dakolias, M. (1999). *Court performance around the world: a comparative perspective*. The World Bank.

EUROSTAT (2016). European statistical office. Retrieved from `https://ec.europa.eu/eurostat`.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the penn world table. *American economic review*, 105(10):3150–82.

ICPSR (2010). United nations surveys of crime trends and operations of criminal justice systems series, waves 1-10, 1970-2006. Retrieved from `https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/26462`.

IMF (2012). Burundi: Poverty reduction strategy paper 2. *IMF Country Report*, 12(224).

Kühn, Z. (2011). *The judiciary in Central and Eastern Europe: mechanical jurisprudence in transformation?* Brill.

Marshall, M. G., Jaggers, K., and Gurr, T. R. (2002). Polity iv dataset. *Computer file*.

Massenot, B. (2012). Why do some countries have so many more lawyers than others? *Working Paper*.

OAS (2016). Organization of american states. Retrieved from http://www.oas.org/dsp/Observatorio/database/indicatorsdetails.aspx?lang=en&indicator=72.

Pistor, K., Wellons, P. A., and Sachs, J. (1999). *The Role of Law and Legal Institutions in Asian Economic Development, 1960-1995*. Oxford University Press, USA.

Schmiegelow, M. and Schmiegelow, H. (2014). *Institutional Competition between Common Law and Civil Law*. Springer.

The World Justice Project (2019). World justice project rule of law index. Retrieved from The World Justice Project, http://worldjusticeproject.org.

Turner, B. (2009). *The Statesman's Yearbook 2010: The Politics, Cultures and Economies of the World*. Springer.

UNODC (2016). United nations office on drugs and crime. Retrieved from https://www.unodc.org/unodc/en/data-and-analysis/statistics.html.

World Bank (2018). World development indicators database. *The World Bank*.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2019/3082/9