The (in)efficiency of Justice. An equilibrium analysis of supply policies

Antonio Peyrache*   Angelo Zago†

Abstract. In this paper we propose an equilibrium computational model of the market for justice that focuses on supply policies aiming to increase the efficiency of the system. We measure performance in terms of completion times and inefficiency in terms of the discrepancy between observed completion time and an efficient benchmark (equilibrium) completion time. By using a rather general production model that can take into account resource use, we can study the (steady state) performance of the justice sector as a whole and improve both on the analysis of length of trials and on standard measures of partial productivity (like the number of defined cases per judge). In order to identify demand and supply and run our counterfactual equilibrium analysis, we focus on a recently collected dataset on the Italian courts of justice system. The Italian case is useful because it provides exogeneous variation in the quantity of interest that allows for identification. It is also interesting because of the heterogeneity of the system in terms of completion times. Overall we find that three supply policies can make a significant contribution to the efficiency of the system: introduction of best practices, break-ups of large courts of justice into smaller ones (to exploit economies of scale), and optimal reallocation of judges across courts (in order to enhance efficiency). We find that, even without introduction of best practices, break-ups and reallocation can reduce the system completion time by around 30%. Although we are critical about the external validity of our results, these results point to the fact that there is large scope for supply policies aiming at improving the processing time of judicial systems.

Keywords: Courts of Justice; Efficiency; Equilibrium; Completion Times.

*School of Economics, University of Queensland, St. Lucia, QLD 4072, Australia. a.peyrache@uq.edu.au

†Dipartimento di Scienze Economiche, Università degli Studi di Verona, Via Cantarane, 24, Verona 37129, Italy. angelo.zago@univr.it

1   Introduction

The judicial system (and the rule of law) has a very important role in securing property rights and enforcing contracts, thus affecting economic behavior, investment choices, and economic growth (Aldashev, 2009). An inefficient judicial sector may negatively impact (to name a few) credit markets (Jappelli et al., 2005; Ponticelli and Alencar, 2016), entrepreneurship (Chemin, 2009), investments (Chemin, 2012), firms size and growth (Kumar et al., 2001; Giacomelli and Menon, 2017), labor markets (Ichino et al., 2003), FDI (Nunn, 2007), public procurement (Coviello et al., 2018) and housing markets (Mora-Sanguinetti, 2012). In addition, as was early recognized by Adam Smith (1776, in Landes and Posner, 1979), the administration of justice is "one of the few proper functions of government", since "private security and enforcement, while working well in some environments, often degenerate into violence" (Djankov et al., 2003, p. 454). The judiciary is therefore one sector of the economy where the market system cannot work properly, given the absence of a functioning output price mechanism that could sanction inefficient units. Moreover, in many countries, citizens cannot "vote with their feet" because cases are assigned where they occur and the involved parties cannot choose the court where to settle them. Thus "it is crucial to understand the factors that make courts function more or less effectively" (Djankov et al., 2003, p. 454), both in developed and developing countries (Palumbo et al., 2013).

In this paper we undertake an equilibrium analysis of the market for justice by focusing on supply policies aiming to increase the efficiency of the system. While most of the studies we are aware of (which will be shortly introduced) consider the demand of justice, we provide a model that can account for inefficiencies arising from its supply side. A measure of performance that practitioners often use is the average length of trials [1] as a way to spot problematic countries or courts within a country (see, e.g., CEPEJ, 2016). However, looking only at completion time without considering resource use is not fully informative and may be misleading. By using a rather general production model that can take into account resource use, we can study the (steady state) performance of the justice sector as a whole and

---

[1]As explained in a recent OECD report, "the focus on length is motivated not only by the importance of a timely resolution of disputes for the correct functioning of the economy, but also by the fact that a reasonable trial length is a necessary (though not a sufficient) condition for good performance in other dimensions [..] Also, as emphasized by the adage *justice delayed is justice denied*, timeliness is a prerequisite for achieving justice. Moreover, the length of trials is also generally associated with other crucial measures of performance such as confidence in the justice system" (Palumbo et al., 2013: p. 9).

improve both on the analysis of length of trials and on standard measures of partial productivity (like the number of defined cases per judge). We can then relate these performance measures (i.e., the time needed to define cases) to the possible causes of long trial length. In particular, we consider three supply policies: introduction of best practices at the court level, break-ups of large courts of justice (to exploit scale economies) and optimal reallocation of judges across courts. We consider the effect of these policies both on the average completion time of the system and on the distribution of completion times for the different courts.

In order to account for resource use, we consider a production frontier for the court of justice where the number of pending cases is a *variable* input and the number of judges is a *fixed* production factor (a capacity input). For a given number of judges, when the number of pending cases increases, the number of defined cases first increases, then reaches a maximum and finally decreases due to congestion effects. For this reason the model can accommodate both variable returns to scale and congestion of production.[2]

The supply policy analysis is conducted taking into account possible demand effects and is therefore investigated in terms of a notion of steady state equilibrium of the system. We suggest that observed slow processing times may be the outcome of a steady state equilibrium with a high number of pending cases (long queues). Implementing the aforementioned supply policies would transition the system to a new steady state equilibrium with low processing times and short queues.

Identification of the demand and supply functions can be problematic due to a possible simultaneity bias (in fact demand and supply is used as a classical example of such a bias). In order to identify and separate demand and supply effects, we make use of a recently collected dataset for the Italian courts of justice system. In Italy, cases are assigned where they occur and we do not observe significant population migration between the different districts in the observed time span. We can use these two conditions to identify the demand function and the elasticity of demand to processing time. This means that we are able to identify demand and supply and then run a counterfactual analysis.

The Italian case is not only interesting because of the aforementioned identification strategy, but it is also interesting due to its heterogeneity [3]. In

---

[2]Congestion may be due, for instance, to *task juggling* (Coviello et al., 2014). The first paper to acknowledge congestion problems in courts is probably Buscaglia and Dakolias (1999), but to the best of our knowledge no other paper has considered it, apart from Dimitrova-Grajzla et al. (2012), Coviello et al. (2015), and Bray et al. (2016).

[3]In the south, for instance, trials take longer and stocks of pending cases are larger. However, given that southern courts are provided with more human resources, it is impor-

fact, Italian courts of justice are among the most inefficient in OECD countries in terms of trial length (Palumbo et al., 2013). According to the World Bank (Doing Business, 2017 edition), Italy ranks 108th out of 183 countries in terms of enforcing contracts, compared to Germany (17th), France (18th), Spain (29th), and UK (31st). The *average disposition time* for a standard commercial case in 2016 was 1,120 days in Italy, against 553 days in OECD countries (regional average), 395 days in France, 499 days in Germany, 510 days in Spain. The average trial's length is quite different across Italian regions as well, with lengthier processes and larger stocks of pending cases in the South of Italy (Carmignani and Giacomelli, 2009).

Given this lacklustre performance, the Italian justice system has been investigated quite extensively, and there has been a lively discussion about the possible causes of these inefficiencies, in particular about the *pathological demand* effects (Marchesi, 2003), according to which higher litigation rates could be explained by long trials. Delays in delivering justice would lead some economic agents (households, workers, and firms) to exploit these inefficiencies to strategically postpone their contractual obligations with other parties. Moreover, these opportunistic choices may be more likely to emerge with increasing differences between legal and market interest rates (see, e.g., Marchesi, 2003; Felli et al., 2008; Padrini et al., 2009). Other theories have suggested *supplier-induced demand* reasons, (see, e.g., Carmignani and Giacomelli, 2010 and Buonanno and Galizzi, 2014), according to which the combination of the increase in the number of lawyers and the minimum fee regulation would lead to excessive litigation. However, the empirical evidence regarding these demand-side possible causes is rather ambiguous, and different studies have been calling for a complementary supply-side analysis (see, e.g., Bianco and Palumbo, 2007; Felli et al., 2008). Indeed, although placed among countries displaying high litigation rates, Italy is given as the example where "there is scope for improvements also on the supply side, for instance expanding the use of case-flow management techniques" (Palumbo et al., 2013: 45), which is a policy aimed at introducing best (management) practices.

To empirically test our model's predictions, we collected data for all Italian courts of justice (165) for the period 2005-2012,[4] taking advantage of data now publicly available and collecting additional data from other sources. Overall, we find that technical (best practices), size (break-ups) and reallocation inefficiencies are the major issues at the industry level. Congestion

---

tant "to establish whether and to what extent the larger stock of pending cases is due to lack of resources or to their lower productivity" (Carmignani and Giacomelli, 2009: 21).

[4]This is before the change in court geography implemented at the end of 2012 by the Monti's Government.

4

inefficiency does not have a big impact at the industry level, but it is a major problem in some specific courts. Although the external validity of our results may be questioned, at least for the Italian case we find that supply policies may give a significant contribution to the overall efficiency of the system in terms of completion times. These results point to the fact that there is large scope for supply policies aiming at improving the processing time of judicial systems in general.

In section 2 we introduce a model of the market for justice. After briefly reviewing some of the major contributions that look at justice, we introduce our model and specification for the supply of justice. In section 3, we introduce the computational models and in section 4 the empirical results. Section 5 concludes with some suggestions for further research.

2 The market for justice

Most of the studies on the "market for justice" consider quite extensively its demand side (see the review in Cooter and Rubinfeld, 1989, for the market for legal disputes and, for instance, Buonanno and Galizzi, 2014 for the market for legal services, i.e., lawyers). In this paper we focus on an equilibrium analysis of the market for justice considering both the demand and supply of justice. Given the absence of a price mechanism, in the market for justice expected completion time plays the role of a rationing mechanism that allows balancing demand and supply in equilibrium.

2.1 The demand for justice

In the market for justice, demand and supply of justice meet. We can measure the demand for justice with the number of incoming cases, and supply by the number of cases defined, i.e., resolved, over the same period. Cooter and Rubinfeld (1989) summarize different contributions of the literature on the market for justice and propose a 'hybrid' model (of previous contributions) that takes into account different actors and different stages of the legal disputes to provide the micro-foundations of the demand for justice. First, in stage 1, a party may act, causing some harm to another party, after having decided on possible precautions to avoid (or reduce the probability of) the harmful event with a standard analysis of costs and benefits. In stage 2, the harmed party has to decide whether to assert the claim or not. Again, her decision depends on a comparison between the costs of asserting the claim, e.g., hiring a lawyer, and the benefits, such as the expected proceeds from the settlement or the victory in the trial. In stage 3, the two parties must decide

on whether to settle the case without going to trial or to go all the way down to court. Last, in stage 4, the case is brought to court where it is discussed and resolved. These different stages are analyzed in a backward fashion. The parties' behavior at trial is based on the merit of the case and on the efforts devoted by both parties. Each party decides her own effort level based on her subjective expected return from that choice (i.e., a possible gain for the plaintiff and a loss for the defendant). Different exogenous variables can affect these effort levels, such as adjustment in compensatory damages, reputation effects, etc. On a similar note, "there is more scope for settlement when litigation is costly [..], negotiations are inexpensive [..], and the disputants are pessimistic about trial outcomes" (Cooter and Rubinfeld, 1989: 1076). From their hybrid model, Cooter and Rubinfeld (1989) then suggest a structural model for empirical research on the trial/settlement split by assuming that the plaintiff's expected gain (defendant's expected loss) from trial consists of a systematic component and a randomly distributed error. From this, one can obtain a reduced-form model in which the probability of trial can be determined by evaluating the probability distribution function of the systematic components empirically. They also survey some papers that empirically implement some variants of such models.

An important dimension in modeling and estimating the demand for justice is trial length. Among the first to recognize it explicitly in the case of courts is probably Gravelle (1990). Economists have long argued that economic efficiency reasons should induce to prefer rationing by price than by waiting (see Barzel, 1974)[5]. However, a trial is not a standard good, because its demand is generated by a sequence of decisions[6], and courts are rationed by waiting lists, not waiting lines. The bottom line is that "it can be efficient to ration by waiting" (Gravelle, 1990: 270). However, given that a trial occurs when an accident happens and the same is not settled out of court, and that trial length affects the probabilities of both events, often in different directions, the overall effect of the expected time to settle a case on the demand for trial may be ambiguous, and thus needs to be empirically estimated.[7]

---

[5]The intuition is that the full price of a good may include both the money price and the opportunity cost of the time waiting. In case of lengthy waiting times, an increase in the money price of the good leads to a lower waiting time. The market clearing full price of the good does not change - and so consumers are equally better off - but the suppliers are better off. Rationing by waiting times is inefficient because it imposes a loss to consumers not compensated by any gain to suppliers.

[6]There are at least two sets of decisions (and relevant affecting variables): the decision to commit a crime (or incur into an accident), and the decision to settle the case out of court.

[7]Jappelli et al. (2005) is among the few papers looking at the effect of trial length on

In effect, different papers have tried to estimate empirically the demand for justice in different countries. Felli et al. (2008) fits into this literature, being probably also the closest to our contribution.[8] They develop an economic model of a litigant's decision to go to trial or to settle following, among others, Van Wijck and VanVelthoven (2000), basing these choices on the expectations regarding the possible outcome of the case, the costs of proceedings and the costs of negotiation. They then suggest an empirical specification in logarithmic first differences, regressing the demand for justice (incoming cases) as a function of different explicative variables, such as the average length of trials, average real earnings for and the number of lawyers, the real market and legal interest rates, some proxies for the business cycle, and a dummy for the year (1995) in which a significant institutional reform was introduced. Estimating their model with a panel of the 26 districts over the period 1991-2002 using fixed effects and GMM, they find that almost all variables are significant (R^2 between 0.64 and 0.8) and, more important for our purposes, that a 1% increase in trial length growth was associated with a demand growth reduction of about 0.75%.

There are other studies that estimated empirically the demand for justice. Ginsburg and Hoetker (2006), for instance, estimate the demand for justice in Japan and test among different possible explanations - cultural, institutional, political ones - for the historical low level and subsequent increase in litigation rates. They take into account the increase in the number of lawyers and judges, the procedural and legal reforms, and the structural changes in the economy. Using data on 47 prefectures for the 1986-2001 period, they find that institutional constraints explain the relatively low rate of litigation in Japan. Indeed, the shortage of lawyers, the shortage of judges, and procedural barriers show results consistent with the lack of institutional capacity in the legal system as a major predictor of low litigation rate. Moreover, the paper rejects the hypothesis that the availability of alternative methods of dispute resolution may be a key factor suppressing litigation, and also that cultural differences may be important. Last, their analysis suggests that absolute levels of wealth increase litigation, but economic decline does it as well. However, they do not consider explicitly whether the time to resolve a case in court can influence the demand for justice or not.

A related literature has tried to shed light on the causes of excessive Italian trial length focusing on the market for lawyers. Buonanno and Galizzi (2014), for instance, suggest a different explanation for the emergence

---

the decision to forfeit a contractual obligation by opportunistic borrowers. They find a significant effect of judicial length on these decisions and thus on the functioning of Italian credit markets.

[8]They also provide a relatively updated survey of this literature.

of long trials, based on the degree of competition in the market for lawyers. The idea is that when the number of lawyers increases and when a minimum fee is imposed (as in the Italian case), lawyers may opportunistically induce their client to go to court more often (than optimal from the client's point of view). They provide empirical evidence to test this supplier-induced demand hypothesis. Since new lawyers may locate where demand is higher, causing reverse causality, they use two instruments for the number of lawyers: a geographic one and a historical one.[9] Their main result is that the number of lawyers operating in a court exerts a positive and statistically significant effect on the litigation rate. This result is robust to different checks, controlling for the general social and economic conditions of different Italian provinces, their urbanization rate and their level of human capital. Last, Buonanno and Galizzi (2014) find that the demand for justice is inversely related to the trial length (they use lagged length) as well.

Carmignani and Giacomelli (2010) perform a similar analysis on Italian data for the years 2000-2005 as well. They document a positive correlation between lawyers and litigation. They also use a 2SLS approach to control for endogeneity, using provinces proximity to a law school in 1975, confirming that the number of lawyers has a quite large positive effect on civil litigation. The results hold with different specifications, and after controlling for economic conditions and economic cycle, social capital, urbanization, and crime rates. Last, including the lagged value of trial length, they find however a positive relationship between the trial length and the demand for justice (as measured by incoming cases).[10]

2.2   The supply of justice

We now switch to the supply side of the market for justice, and assume that the supply of justice is related to the resources that are used, given the available technology and the efficiency in its use. Moreover, the supply may depend as well on management, legal formalism (Djankov et al., 2003), incentives for the judges, etc. (see, e.g., Palumbo et al., 2013: 9). In general, we can specify these relationships and the technology with the following

---

[9]Notice that Ginsburg and Hoetker (2006) as well use an instrumental variable approach for the number of lawyers and judges.

[10]Mora-Sanguinetti and Garoupa (2015) undertake a similar study for Spain, a country in which the litigation rate is even higher than Italy (Palumbo et al., 2013). They use a similar IV approach, finding no clear evidence of endogeneity, but a clear positive relationship between the number of lawyers and the induced demand for justice. They do not consider the effect of trial length on demand though.

production function

$$y = \theta F(g, p) \tag{1}$$

where the 'production' of defined cases y, depends on the number of judges g and the number of pending cases p; and the parameter $\theta \in [0, 1]$ represents the efficiency of production. We assume that the shape of the production function conforms to the law of variable proportions as described in Svensson and Färe (1980): the number of judges is considered as a capacity factor which is processing the number of pending cases (the variable factor of production). For a given number of judges, when the number of pending cases increases, the number of defined cases first increases, then reaches a maximum (the full capacity) and then it declines due to congestion effects. This is represented in figure 1, with the number of pending cases on the x-axis and the number of defined cases on the y-axis. The curve in figure 1 depicts the production possibility set for a given number of judges.

Conditioning the analysis on the number of judges is important for at least two reasons. First, there can be significant differences between small courts and bigger ones.[11] Second, previous studies have found different returns to scale in Italian courts.[12] We also need to consider congestion, as already documented in other studies.[13] Indeed, "the inability of the system to satisfy the demand for justice (i.e., resolve in each given period a number of cases equal to that brought to court) generates congestion and delays" (Palumbo et al., 2013: 10).

Our production function may appear like in Figure 1. Notice that up to point 1, there are increasing returns, while between 1 and 2 the returns are decreasing. At point 2, the system reaches full capacity. On the other hand, beyond 2 congestion starts to kick in. The slope of the CRS closure line is $y/p = 1/t$, that is the inverse of the time needed to process a case with that capacity and given the flow of incoming cases. Processing time is therefore defined as the ratio of pending cases to the number of defined cases. This is a standard measure in queueing theory and it has a very appealing simple interpretation. Last, notice that a change in 'capacity', that is in the fixed factor of production(s), would shift the production frontier and potentially enlarge the production possibilities of the courts of justice.

While in our analysis it would be difficult to consider legal formalism,

---

[11]For instance, in 2012 the biggest court in Italy (Rome) had 406 judges, while the smallest ones (5 courts) had 6 judges.

[12]Marchesi (2003), for instance, found that courts with fewer than 20 judges were too small, while courts with more than 80 judges suffered from decreasing returns to scale.

[13]Coviello et al. (2015), for instance, for the court of Milano, found that "if a larger future caseload induces judges to increase task juggling by 1%, [..] the completion hazard would decline approximately by a factor ranging between 1% and 2.1%" (p. 909).
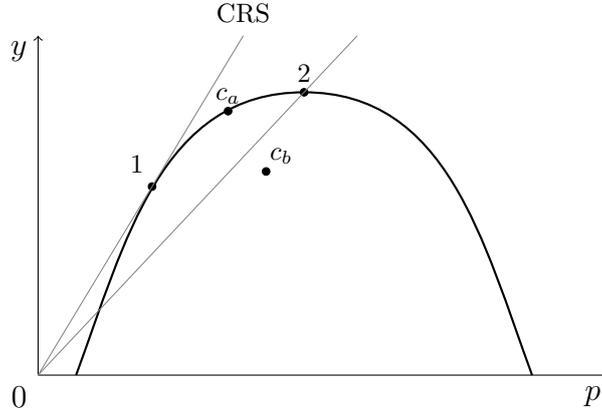
Figure 1: The production of justice

judges' incentives and other determinants that are common across all Italian courts, we can explicitly take into consideration their different efficiency levels. We in fact consider this production model within the framework of efficiency analysis, so that a specific court can be located on the production frontier ($\theta = 1$), for example court c_a, or inside the frontier like court c_b, where the distance to the frontier represents a measure of inefficiency. In this fashion, we can measure the inefficiencies for each court and for the overall industry following Peyrache and Zago (2016).

An alternative way of representing this production trade-offs is in the time-production (t, y) space rather than the just introduced pending-production space (p,y), where t is trial length and y are the defined cases. Given the assumptions of our production model, we obtain the supply function S depicted in Figure 2. Notice that the backward bending portion of the supply function here represented comes from the congestion hypothesis.[14] Again, a change in the number of judges will shift the supply function in the time-production space as well as the production function in the production-pending space.

2.3 Equilibrium

The equilibrium between demand and supply is obtained by considering what we will refer to as a material balance condition. This condition states that the variation in the number of pending cases will be equal to the discrepancy between the number of incoming cases (inc) and the number of defined cases each year:

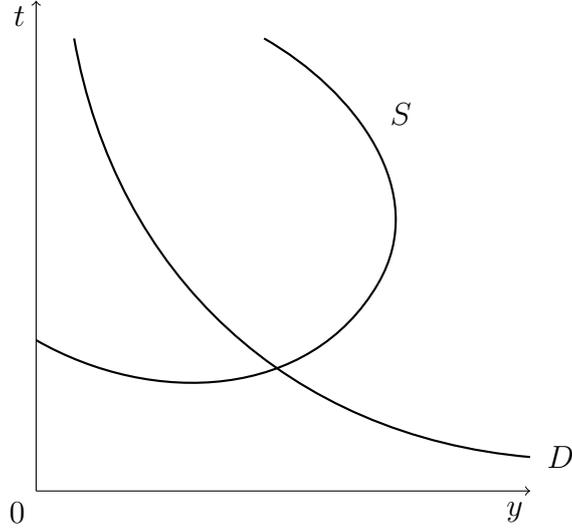[14]In the technology we allow for this possibility, but we test its relevance with our data in the next section.

Figure 2: The equilibrium in the market for justice

$$p_{t+1} - p_t = inc_t - y_t. \tag{2}$$

We say that the system is in steady state if the number of pending cases stays constant from one year to the next: $p_{t+1} = p_t$. Or, to state the same in a different way, the system is in steady state if the number of incoming cases is equal to the number of defined cases each year.

Since (differently from normal markets that are cleared by prices) the market for justice clears through variations in the length of proceedings (completion time, see Palumbo et al., 2013: p. 9), the equilibrium conditions for the market for justice will imply a steady state condition

$$inc_t = y_t, \tag{3}$$

and a standard equilibrium condition in terms of demand and supply $S(t) = D(t)$. This equilibrium will provide a completion time which will generate a demand exactly equal to the number of defined cases. We say that the system is in a steady state equilibrium if these conditions hold, and we use this notion of steady state equilibrium to assess counterfactual changes in the production efficiency of the courts of justice.

In figure 3 we represent the same equilibrium conditions in the pending cases-defined cases space (which is more natural when thinking about production). We can invert the demand function in order to obtain the combinations of pending-defined cases that are compatible with the demand conditions. In other words the demand line in this space determines all the combinations of
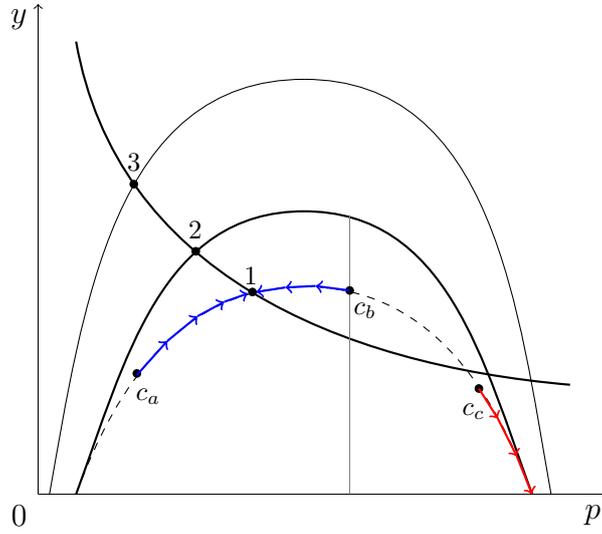
11

Figure 3: Efficiency and equilibrium

pending and defined cases that will make the number of incoming cases (as determined by the demand function) equal to the number of defined cases.

In this figure we distinguish a number of effects. First, there are three reference production functions. The one passing through point 1 is the production that can be achieved for any given quantity of pending cases with a given level of efficiency $\theta < 1$. The steady state equilibrium for this level of efficiency is at point 1 where demand and supply meet. If a court is at a point like c_a, then the number of incoming cases will be larger than the number of defined cases and the number of pending cases will increase, moving the court towards the equilibrium point. For points like c_b the number of defined cases is larger than the number of incoming cases, thus the number of pending cases will decrease, moving this point towards the equilibrium. This means that equilibrium 1 is a stable steady state equilibrium. There are then points like c_c with a number of incoming cases which is larger than the number of defined cases: this will have the effect of increasing the number of pending cases with the consequence of congesting production even more. This last set of points will converge to a point of full congestion. In the next section we identify all these three types of courts.

One last thing that we notice in this picture is that if one starts at the equilibrium point 1, an increase in production efficiency will have the effect of increasing the number of defined cases above the number of incoming cases, reducing the stock of pending cases and moving the system towards a new steady state equilibrium depicted by point 2. At this new equilibrium the

12

court of justice will be fully efficient ($\theta = 1$) and it will be in steady state equilibrium. Similarly, an increase in production capacity, coming either from an increase in the number of judges or a better exploitation of scale economies, will have the effect of shifting this frontier even further towards a new steady state equilibrium 3.

This is our basic conceptual framework to analyze improvements in the efficiency of the system coming from supply side policies aiming at introducing best practices, implementing break-ups and reallocating judges efficiently across the different courts of justice.

3 Methodology

The first step in our procedure is to estimate the demand for justice. To this purpose we consider the following regression model:

$$\log inc_{it} = \alpha_i + \mathbf{X}_{it}\boldsymbol{\beta} - \gamma \log t_{it-1},$$

where $i = 1, \ldots, n$ indexes courts of justice, $t = 1, \ldots, T$ the reference year. We have a panel of 165 courts of justice (the whole population of courts in Italy) for the years 2005-2012. The variable inc_{it} is the number of incoming cases in year t and court i and $\mathbf{X}_{it}$ contains control variables, including individual level dummy variables and the population size of each court of justice district (this is exogenously determined, since each population district makes reference to a given court of justice and movements between courts is forbidden by law). The average completion time for a new case is given by the ratio of the number of pending cases to the number of processed cases $t_{it} = p_{it}/y_{it}$. Since we also have data on the number of judges (g_{it}) and the number of defined (y_{it}) and pending cases (p_{it}), we can compute the queue average completion time for each court in each time period. In the demand equation we include the lagged value of the processed time. This can be interpreted both as a causal relationship of the completion time onto the number of incoming cases (quantity demanded) under the assumption of adaptive expectations for the plaintiff, or as a prediction equation for the number of incoming cases. To check the robustness of our estimates, we also tested the sign and size of demand elasticity (γ) by using the lagged value of the number of incoming cases as an instrument for the following contemporaneous demand equation (based on rational expectations):

$$\log inc_{it} = \alpha_i + \mathbf{X}_{it}\boldsymbol{\beta} - \gamma \log t_{it}.$$

The sign and size differences of the γ coefficient are negligible, therefore we use the first equation as it has a simpler interpretation. The steady state

equilibrium condition specifies that the number of processed cases each year must be equal to the number of incoming cases each year:

$$inc_{it} = y_{it}.$$

This equilibrium condition, together with the material balance condition (specified in the previous section) implies that in equilibrium the queue at the court of justice is in steady state with the number of pending cases constant from one year to the next, which in turns implies (unless some of the control variables have changed from one year to the next) that $inc_{it} = inc_{it-1}$. The equilibrium condition and material balance condition imply that we can derive a relationship between the number of pending cases and the number of defined cases from the demand function:

$$y = p^{-\frac{\gamma}{1-\gamma}} \exp\left(\mathbf{X}\boldsymbol{\beta}/\left(1-\gamma\right)\right).$$

This last equation is useful in order to analyze demand trade-offs in the pending-defined cases space. This also implies the following equilibrium relationship between the average completion time and the number of pending cases (dividing the previous equation by p and taking the inverse):

$$t = \frac{p^{\frac{1}{1-\gamma}}}{\exp\left(\mathbf{X}\boldsymbol{\beta}/\left(1-\gamma\right)\right)},$$

where $\mathbf{X}\boldsymbol{\beta}$ is the prediction based on the demand regression estimates.

We note that, unless demand is rigid with respect to completion time ($\gamma = 0$), increasing efficiency of production (by increasing the number of processed cases) will decrease completion time and increase the number of incoming cases. Therefore efficiency gains may be overestimated if the demand side is ignored. We find demand to be quite inelastic at a value of $\gamma = 0.16$. This means that in our dataset a 10% increase in completion time will reduce the number of incoming cases by 1.5%.

We now turn our attention to the supply side of the "market" for justice where we incorporate the equilibrium condition in the determination of an efficient configuration of the system. This will return the *equilibrium* increase in the efficiency of the system rather than a simple *optimal* increase condition (which could be out of equilibrium). Clearly, an equilibrium increase in efficiency will also be optimal, but the reverse is not in general true. For example, in the analysis of Peyrache and Zago (2016) the efficiency of the system is studied from an optimality perspective, by looking at the potential increase in output given the number of incoming cases and the number of pending cases. This may result in an overestimation of the potential efficiency

gains once the equilibrium condition is introduced (i.e., the output oriented efficiency measure is not an equilibrium notion). Since we are looking at equilibrium outcomes, in the following analysis we take the individual average of each variable in the panel, and consider the following data structure:

$$(\mathbf{G}, \mathbf{Y}, \mathbf{P}, \mathbf{X}\boldsymbol{\beta})$$

where $\mathbf{G}$ is the vector with the average number of judges ($g_i = \frac{1}{T}\sum_t g_{it}$) for each court, $\mathbf{Y}$ is the vector with the average number of defined cases ($y_i = \frac{1}{T}\sum_t y_{it}$), $\mathbf{P}$ is the vector with the average number of pending cases ($p_i = \frac{1}{T}\sum_t p_{it}$) and $\mathbf{X}\boldsymbol{\beta}$ is the vector of predicted incoming cases from the demand regression.

We start with a linear program that defines the level of efficiency of court of justice i, which is processing y_i cases with g_i judges and is facing p_i pending cases (we always denote decision variables with greek letters):

$$\min_{\lambda_{ik}} \ \theta_i$$
$$st \ \sum_k \lambda_{ik} g_k \leq g_i$$
$$\sum_k \lambda_{ik} p_k = p_i \qquad (4)$$
$$\sum_k \lambda_{ik} y_k \geq y_i/\theta$$
$$\sum_k \lambda_{ik} = 1.$$

This is known as the data envelopment analysis (DEA) estimator of efficiency (see Coelli et al., 2005). The only difference with respect to a standard DEA model is that we impose an equality constraint for the number of pending cases. This equality constraint, together with the standard inequality constraints associated with the number of judges and the number of defined cases, allows for the law of variables proportions depicted in figure 1: given the number of judges (our capacity measure) when the number of pending cases increases, the number of defined cases will first increase, reach a maximum and then decrease (due to congestion effects). The efficiency score θ_i is measuring the distance from the frontier in terms of additional number of pending cases that could be processed when the court of justice is benchmarked against other courts of justice of similar size. This is therefore a measure of best practice efficiency.

We notice that neither the observed combination (y_i, p_i, g_i) nor the efficient combination $(y_i/\theta, p_i, g_i)$ are necessarily equilibrium levels of produc-

15

tion, since the processed time induced by these quantities may be incompatible with our demand equation. In order to define an equilibrium outcome that keeps the efficiency level of the court of justice constant at the observed level θ_i we pose the following program where we explicitly include the relationship derived from the demand function:

$$\min_{\lambda_{ik},\pi_i,\tau_i} \quad t_i = \frac{\pi_i}{\tau_i}$$

$$st \quad \sum_k \lambda_{ik} g_k \leq g_i$$

$$\sum_k \lambda_{ik} p_k = \pi_i$$

$$\theta_i \sum_k \lambda_{ik} y_k \geq \tau_i$$

$$\sum_k \lambda_{ik} = 1$$

$$\tau_i = \pi_i^{-\frac{\gamma}{1-\gamma}} \exp\left(\mathbf{X}_i \boldsymbol{\beta} / (1-\gamma)\right).$$

In this program the number of pending cases π_i and the number of defined cases τ_i are decision variables and they are chosen in order to minimize the average completion time of the court of justice $t_i = \pi_i/\tau_i$. The last constraint in this optimization program is the demand constraint as derived at the beginning of this section. Since the level of efficiency θ_i is kept constant at the level determined with program (4), the completion time as determined by this program is an equilibrium completion time for the given level of efficiency θ_i. In other words, this program returns the equilibrium average completion time for the efficiency level θ_i as opposed to the observed non-equilibrium level $t = p_i/y_i$. We notice that in some extreme cases if the court is operating in the congestion part of the production frontier, then the number of pending cases will grow unbounded and the equilibrium will be $\tau = 0; \pi = max$. To detect these cases it suffices to check that $inc > y$ and $p > \pi$. In all other cases the equilibrium time is well defined and the system will eventually converge to the optimal time. This optimal time can be lower or higher than the observed time depending on the starting point and the shape of the estimated production frontier. For courts that are heading towards full congestion the only solution is to increase resources (the number of judges) or increase efficiency of production. We discuss this issue in the following section on results.

The optimization program just introduced is non-linear, but it is easy to show that it is convex, by substituting the demand equation into the third

constraint and into the objective function, one obtains:

$$\min_{\lambda_{ik}, \pi_i} \pi_i$$

$$st \quad \sum_k \lambda_{ik} g_k \leq g_i$$

$$\sum_k \lambda_{ik} p_k = \pi_i \tag{5}$$

$$\theta_i \sum_k \lambda_{ik} y_k \geq \pi_i^{-\frac{\gamma}{1-\gamma}} \exp\left(\mathbf{X_i}\boldsymbol{\beta}/\left(1-\gamma\right)\right)$$

$$\sum_k \lambda_{ik} = S.$$

$$S = 1 \tag{6}$$

One can readily verify that the function on the right hand side of the third constraint is convex, thus the program can be solved using standard convex optimization solvers[15]. We also notice that we substituted the objective function

$$t\left(\theta_i, S = 1\right) = \frac{\pi_i^{\frac{1}{1-\gamma}}}{\exp\left(\mathbf{X}_i\boldsymbol{\beta}/\left(1-\gamma\right)\right)},$$

with π_i because completion time is monotonically increasing in π_i for $0 \leq \gamma < 1$ (therefore this will have no effect on the optimal solution). We call the equilibrium completion time $t\left(\theta_i, S = 1\right)$ to emphasize that it depends on the current level of efficiency of production. The role of variable S (which is constrained so far to be equal to one) will be clear in the next paragraph.

We are now in a position to ask what happens to the equilibrium completion time, should we introduce best practices and increase the level of efficiency to $\theta_i = 1$. This requires solving the same optimization program (5) with the full efficiency values rather than the observed efficiency values. The comparison of the equilibrium completion time at full efficiency vs the equilibrium completion time at a given level of efficiency will return an equilibrium measure of the efficiency of the court of justice:

$$TE_i = \frac{t\left(\theta_i = 1, S = 1\right)}{t\left(\theta_i, S = 1\right)}.$$

---

[15]We use CVX in Matlab, a package for specifying and solving convex programs; CVX Research, Inc. CVX: Matlab software for disciplined convex programming, version 2.0. http://cvxr.com/cvx, April 2011. M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs, Recent Advances in Learning and Control (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, pages 95-110, Lecture Notes in Control and Information Sciences, Springer, 2008. http://stanford.edu/~boyd/graph_dcp.html.

One should notice that this notion of efficiency is an equilibrium notion, since it includes a constraint that takes into account the behavior of demand when average completion time changes. On the contrary θ_i is a measure of optimality without regards to the level of demand for the service (i.e., it is a potential rather than something that can be realized) and should be used to assess if a court is on the frontier or on the interior of the set. It should also be noted that in general the observed completion time is different from the equilibrium completion time (at the given level of efficiency θ_i) since the observed one may be not respecting our equilibrium notion. Thus equilibrium efficiency is the one that will prevail after all adjustments on the number of pending cases are made and the steady state conditions hold. In other words our measure of time efficiency is comparing two alternative steady states: one with the observed level of inefficiency and the other one with the court of justice lying onto the frontier.

The optimal size of courts of justice. Technical efficiency is far from being the only component of inefficiency in the system. There are other two types that we are going to explore in this and the next section: inefficiencies deriving from scale economies and inefficiency arising from a non-optimal allocation of judges across the different courts.

The most convincing argument in terms of the size of courts of justice is given by looking at the following scatter plot, where we report the number of processed cases and the number of judges. This is consistent with the framework of Fare and Svenson (1980) where one considers the maximal capacity of each court. Scale economies are then derived from the production function with the variable factor at the optimal proportion. Clearly, in the reported scatter, smaller size courts have a better ability at processing cases than larger size courts. In fact production is almost proportional in the number of judges until a size of between 50 and 100 judges and then it declines sharply. This means that a method for increasing efficiency in production is to split larger courts of justice into smaller units. For example, the largest court of justice is obs=118 (Rome) with more than 400 judges: by implementing a break-up policy into subunits the production possibilities of this court could increase dramatically (in fact we found almost a doubling of the production ability for the court in Rome).

In order to quantify the optimal number of splitting for each court of justice, we consider the following capacity model (which is based on Maindiratta, 1990; see also Peyrache and Zago (2016) for a discussion of this effect in the justice sector):
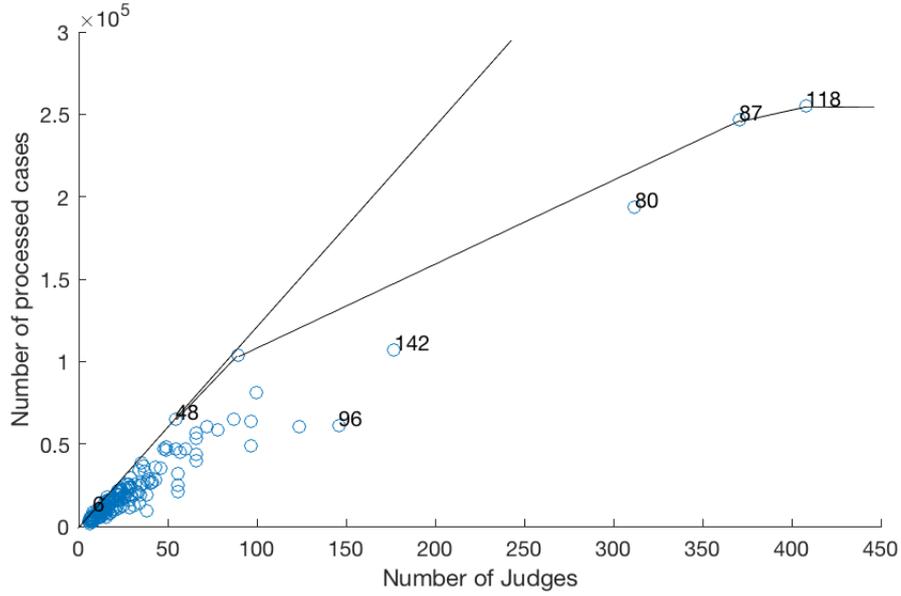
18

Figure 4: Size and efficiency

$$\min \ \beta_i$$

$$st \ \sum_k \lambda_{ik} g_k \leq g_i$$

$$\sum_k \lambda_{ik} p_k = p_i$$

$$\sum_k \lambda_{ik} y_k \geq y_i/\beta$$

$$\sum_k \lambda_{ik} = S_i$$

$$S_i \in \{1, 2, \ldots\}.$$

The optimal values from this program can be used to determine the equilibrium completion time of each court of justice under the scenario of an efficient break-up of large units into an optimal number of smaller units. The optimal values of the integer variable S_i^* will provide the optimal number of sub-units into which large courts should be split. Given these optimal values, we can assess the steady state equilibrium completion time for each court of justice by using these values in program (5). Since this is a relaxation of the original problem, the completion time will be shorter or equal

than the previous optimal one; thus we can pose:

$$SE_i = \frac{t\left(\theta_i, S_i^*\right)}{t\left(\theta_i, S_i = 1\right)}.$$

This is the optimal equilibrium size of the court of justice, accounting for the effect of the demand of justice and keeping the level of efficiency at the current level. It is also possible to consider a joint policy of introduction of best practices and a policy of optimal break-up, obtaining a total time efficiency gain:

$$SE_i = \frac{t\left(\theta_i = 1, S_i^*\right)}{t\left(\theta_i, S_i = 1\right)}.$$

Reallocation inefficiency. The last supply policy we consider is the reallocation of judges between courts of justice. This may enhance the efficiency of the system by moving judges from courts which have very fast processing times to courts with much slower processing times; or, similarly, it could consider gains obtainable when the number of judges are found not to constraint production (i.e., there is large excess capacity). The reallocation problem for the industry can be written in the following way, where we solve in one stage for all the courts of justice:

$$\min_{\lambda_{ik}, \pi_i, \mu_i} \sum_i \pi_i$$

$$st \ \sum_k \lambda_{ik} g_k \leq \mu_i \ , \ \forall i$$

$$\sum_k \lambda_{ik} p_k = \pi_i \ , \ \forall i$$

$$\theta_i \sum_k \lambda_{ik} y_k \geq \pi_i^{-\frac{\gamma}{1-\gamma}} \exp\left(\mathbf{X}_i \boldsymbol{\beta} / \left(1 - \gamma\right)\right) \ , \ \forall i$$

$$\sum_k \lambda_{ik} = S_i \ , \ \forall i$$

$$\sum_i \mu_i \leq \sum_i g_i$$

$$t_i = \frac{1}{\exp\left(\mathbf{X}_i \boldsymbol{\beta} / \left(1 - \gamma\right)\right)} \pi_i^{\frac{1}{1-\gamma}} \leq t_{max}.$$

This problem is looking at the minimization of the system completion time, given demand constraints for each court of justice, by reallocating judges across courts of justice. The optimal value of the decision variables μ_{it} will provide the optimal reallocation of judges in order to minimize the

	Current efficiency - No break-ups	Full efficiency - No break-ups	Current efficiency - Break-ups	Full efficiency - Break-ups
Without reallocation of judges	1	2	3	4
With reallocation of judges	5	6	7	8

Table 1: Possible supply policy scenarios

average completion time of the system. It should be noted that by setting $\mu_i = g_i$, one obtains the individual court efficiency problem (5) in stacked form (and with identical solution). Therefore the reallocation of judges cannot be pejorative of the current configuration. In this program there are two parameters that can be changed: θ_i and S_i. These parameters are used to determine alternative supply policy scenarios in terms of best practices and break-up policies. In particular it is possible to look at 4 combinations: current efficiency θ_i and no break-ups $S_i = 1$; current efficiency with break-ups S_i; full efficiency without break-ups and full efficiency with break-ups.

4 Empirical results

For the alternative 8 scenarios considered in table 1, we report both the overall equilibrium completion time for the system as well as the distribution of completion times for each single court of justice.

In table 2 we report the classification of courts of justice based on the discussion conducted with respect to figure 3. There are three types of courts: type 1 is converging to the equilibrium steady state from the left; type 2 is converging to the steady state equilibrium from the right; and type 3 is diverging from the steady state. We emphasize that type 3 courts are diverging towards a fully congested outcome, yet there exists a steady state equilibrium for these courts. The main problem with these types of courts is therefore that they are unlikely to converge towards the steady state. The table reports the frequency of the three types of courts we observe in the sample. Type 1 represents 51.5% of the total population and it has an average completion time of 10 months and in equilibrium would converge towards a processing time of 11.4 months. Type 2 represents 43.6% of the population and it has an observed completion time of 17.7 months, which is converging

Type	Observed Time	Equilibrium Time	Frequency
1	10.1	11.4	85
2	17.7	14.6	72
3	17.4	8.9	8

Table 2: Observed and equilibrium simple average completion times (in months) for the three types

towards a steady state equilibrium completion time of 14.6 months. Type 3 courts represent only 4.9% of the population; the observed average processing time is 17.4 months and the equilibrium completion time would be 8.9 months, though, as explained before, without further intervention these 8 courts of justice would fail to reach such a fast processing time. Figure 5 shows how the different types of courts are distributed: while types 3 are evenly distributed across Italy, type 2 courts are prevalent in the Southern regions.

Table 3 reports the current processing time for the system as a whole (calculated as the total number of pending cases in the system over the total number of defined cases in the system; this is nothing more than a weighted average of the observed completion times of the individual courts) and the processing time associated with each single supply policy scenario. The current processing time for the system is 14.1 months and the equilibrium completion time (after the system adjusts to the steady state) is 13 months. This is a weighted average for the whole system and in order to grasp the dispersion of processing times across the different courts of justice we report a boxplot of the processing times in figure (6a). From this figure it is clear that the equilibrium completion times for some of the slowest processing courts is diminishing when moving to the equilibrium. The problem with many of the slow processing courts is that they are not necessarily converging to this steady state outcome. These differences are reported in table 2.

Looking again at table 3, one can assess the equilibrium outcome of the different policy scenarios. For example, a policy of break-ups of large courts together with an optimal reallocation of judges (policy scenario 7) would reduce the processing time of the system from 13 months to 9.4 months. This implies only the optimal use of scale economies and the optimal allocation of judges. In particular one can see from the boxplot in Figure (6b) that this policy scenario (number 7) would have quite a substantial effect on the whole distribution of processing times and especially on those courts of justice with the slowest processing times. The slowest equilibrium processing time in the system in scenario 1 (the status quo without any intervention) is
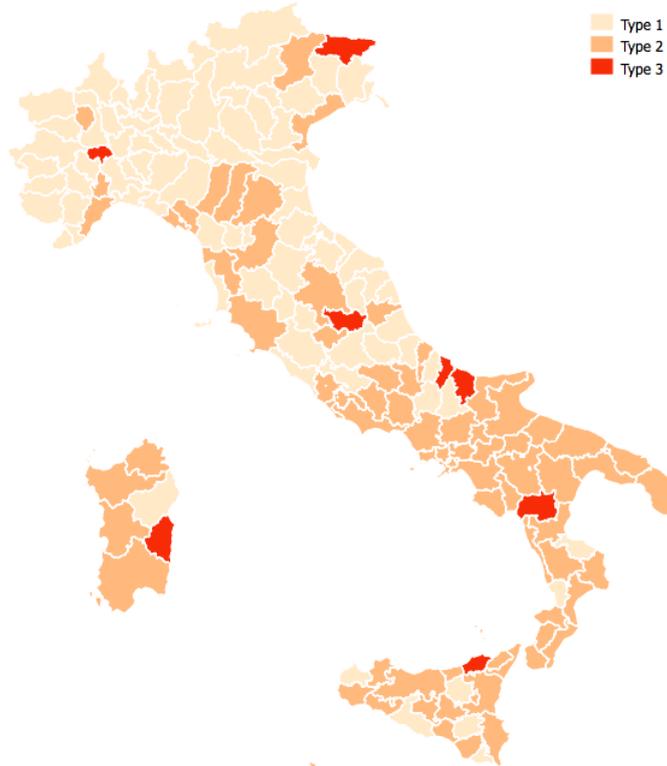
Figure 5: Court classification

around 27 months and this is reduced in scenario 7 to 17 months. This is a substantial reduction considering that this would be the reduction in the slowest processing time in the system.

From table 3 it is also clear that the introduction of best practices has a major effect in reducing completion times in the system. In particular the implementation of policy scenario 8 (introduction of best practices, reallocation of judges and break-ups of the large courts) would bring the system completion time to 6.3 months and would reduce enormously the dispersion in the system with only 1 court of justice taking more than 8 months to process a case. In Figure (7) and (8) we also report the reduction in processing times across different courts under the different policy scenarios. Full efficiency appears the scenario with most of the impact.

Observed processing time is 14.1 months	Current efficiency - No break-ups	Full efficiency - No break-ups	Current efficiency - Break-ups	Full efficiency - Break-ups
Without reallocation of judges	13.0	8.6	11.1	7.2
With reallocation of judges	12.0	7.3	9.4	6.3

Table 3: System time efficiency gains for alternative supply policy scenarios (in months)

5 Conclusion

In this paper we review relevant literature on the efficiency of the justice system and we propose a model to assess the equilibrium efficiency of the justice market. We consider supply policy scenarios based on break-ups of large courts of justice, reallocation of judges across courts and introduction of best practices. The counterfactual analysis takes into consideration that there is a feedback effect from the demand side of the market as a response to the shortening of processing times when these supply policies are implemented. We consider implementation of the different supply policy scenarios in any possible combination and show how the average processing time of the system (and its distribution) varies in these counterfactual analyses. In particular, we find that if all of these supply policies were to be implemented the average completion time of the system would be halved, even accounting for the additional increased demand for justice that this faster processing time would induce. We therefore conclude that the three policy scenarios are sufficient to bring the system down to a processing time which is comparable to other OECD countries.

In this paper we do not consider at least three other issues. The first one has to do with a possible supply policy scenario where the number of judges would be increased (so increasing the capacity of the system) and distributed optimally in order to reduce processing times (without any other supply policy). This would create an additional supply policy scenario in which the resource cost of the system would be increased in order to obtain the desired level of processing time. The second issue we do not consider is that an alternative scenario would be some type of demand policy in which

one may reduce demand by, say, 10% and see what happens to equilibrium completion times. Finally, one issue we do not address is the transition towards the steady state equilibrium. For example when we say that the system is converging from the observed completion time of 14.1 months to an equilibrium completion time of 13 months, we do not specify the timing of this adjustment and it would be interesting in a future research to explore such transitional dynamics.

References

Aldashev, G. (2009), 'Legal institutions, political economy, and development', *Oxford Review of Economic Policy* **25 (2)**, 257–270.

Barzel, Y. (1974), 'A theory of rationing by waiting', *Journal of Law and Economics* **17**, 73–95.

Bianco, M. and G. Palumbo (2007), Italian Civil Justice's Inefficiencies: a Supply Side Explanation, Mimeo, Banca d'Italia.

Bray, R.L., D. Coviello, A. Ichino and N. Persico (2016), 'Multitasking, Multiarmed Bandits, and the Italian Judiciary', *Manufacturing and Service Operations Management* **18(4)**, 545–558.

Buonanno, P. and M. M. Galizzi (2014), 'Advocatus, et non Latro?: Testing the Excess of Litigation in the Italian Courts of Justice', *Review of Law and Economics* **10(3)**, 285–322.

Buscaglia, E. and M. Dakolias (1999), Comparative International Study of Court Performance Indicators, Washington, D.C., The World Bank.

Carmignani, A. and S. Giacomelli (2009), La giustizia civile in italia: i divari territoriali, Questioni di economia e finanza (occasional papers) no. 40, Banca d'Italia.

Carmignani, A. and S. Giacomelli (2010), Too many lawyers? Litigation in Italian civil courts, Temi di discussione (working papers) no. 745, Banca d'Italia.

CEPEJ (2016), European judicial systems efficiency and quality of justices (edition 2016), Cepej studies no. 23, European Commission for the Efficiency of Justice, Strasbourg.

Chemin, M. (2009), 'The impact of the judiciary on entrepreneurship: Evaluation of Pakistan's 'Access to Justice Programme'', *Journal of Public Economics* **93**, 114–125.

Chemin, M. (2012), 'Does court speed shape economic activity? evidence from a court reform in india', *Journal of Law, Economics, and Organisation* **28(3)**, 460–485.

Cooter, R. D. and D. L. Rubinfeld (1989), 'Economic Analysis of Legal Disputes and Their Resolution', *Journal of Economic Literature* **27(3)**, 1067–1097.

Coviello, D., A. Ichino and N. Persico (2014), 'Time Allocation and Task Juggling', *American Economic Review* **104(2)**, 609–623.

Coviello, D., A. Ichino and N. Persico (2015), 'The Inefficiency of Worker Time Use', *Journal of the European Economic Association* **13(5)**, 906–947.

Coviello, D., L. Moretti, G. Spagnolo and P. Valbonesi (2018), 'Court Efficiency and Procurement Performance', *Scandinavian Journal of Economics* **120(3)**, 826–858.

Dimitrova-Grajzla, V., P. Grajzl, J. Sustersic and K. Zajc (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**, 19–29.

Djankov, S., R. La Porta, F. Lopez-deSilanes and A. Shleifer (2003), 'Courts', *Quarterly Journal of Economics* **118(2)**, 453–517.

Felli, E. L., D. A. London-Bedoya, N. Solferino and G. Tria (2008), The "Demand for Justice" in Italy: Civil Litigation and the Judicial System, *in* P.F. and R.Ricciuti, eds, 'Italian Institutional Reforms: A Public Choice Perspective', Springer, New York, NY, pp. 155–177.

Giacomelli, S. and C. Menon (2017), 'Does weak contract enforcement affect firm size? evidence from the neighbour's court', *Journal of Economic Geography* **17(6)**, 1251–1282.

Ginsburg, T. and G. Hoetker (2006), 'The unreluctant litigant? an empirical analysis of japan's turn to litigation', *The Journal of Legal Studies* **35(1)**, 31–59.

Gravelle, H. S. E. (1990), 'Rationing trials by waiting: Welfare implications', *International Review of Law and Economics* **10**, 255–270.

Ichino, A., M. Polo and E. Rettore (2003), 'Are Judges Biased by Labor Market Conditions?', *European Economic Review* **47(5)**, 913–944.

Jappelli, T., M. Pagano and M. Bianco (2005), 'Courts and Banks: Effects of Judicial Enforcement on Credit Markets', *Journal of Money, Credit and Banking* **37(2)**, 223–244.

Kumar, B., R. Rajan and L. Zingales (2001), What determines firm's size, Nber working paper 7208, NBER.

Landes, W. M. and R. A. Posner (1979), 'Adjudication as a private good', *The Journal of Legal Studies* **8(2)**, 235–284.

Marchesi, D. (2003), *Litiganti, avvocati e magistrati. Diritto ed economia del processo civile*, Il Mulino, Bologna.

Mora-Sanguinetti, J. S. (2012), 'Is judicial inefficacy increasing the weight of the house property market in spain? evidence at the local level', *SERIEs* **3**, 339–365.

Mora-Sanguinetti, J. S. and N. Garoupa (2015), 'Do lawyers induce litigation? evidence from spain, 2001–2010', *International Review of Law and Economics* **44**, 29–41.

Nunn, N. (2007), 'Relationship-specificity, incomplete contracts and the pattern of trade', *The Quarterly Journal of Economics* **122(2)**, 569–600.

Padrini, F., D. Guerrera and D. Malvolti (2009), La congestione della giustizia civile in Italia: Cause ed implicazioni per il sistema economico, Note tematiche no. 08, Ministero dell'Economia e delle Finanze.

Palumbo, G., G. Giupponi, L. Nunziata and J.S. Mora Sanguinetti (2013), The economics of civil justice: New cross-country data and empirics, OECD Economics Department working papers no. 1060, OECD.

Peyrache, A. and A. Zago (2016), 'Large courts, small justice! the inefficiency and the optimal structure of the italian justice sector', *Omega* **64**, 42–56.

Ponticelli, J. and L. Alencar (2016), 'Court Enforcement, Bank Loans, and Firm Investment: Evidence from a Bankruptcy Reform in Brazil', *Quarterly Journal of Economics* **131**, 1365–1413.

Santos, S.P. and C.A.F. Amado (n.d.), 'On the need for reform of the Portuguese judicial system - Does Data Envelopment Analysis support it?, journal = Omega, year = 2014, volume = 47, pages = 1–16,'.

Svensson, L. and R. Färe (1980), 'Congestion of production factors', *Econometrica* **48(7)**, 1745–1753.

Van Wijck, P. and B. VanVelthoven (2000), 'An economic analysis of the american and the continental rule for allocating legal costs', *European Journal of Law and Economics* **9(2)**, 115–125.
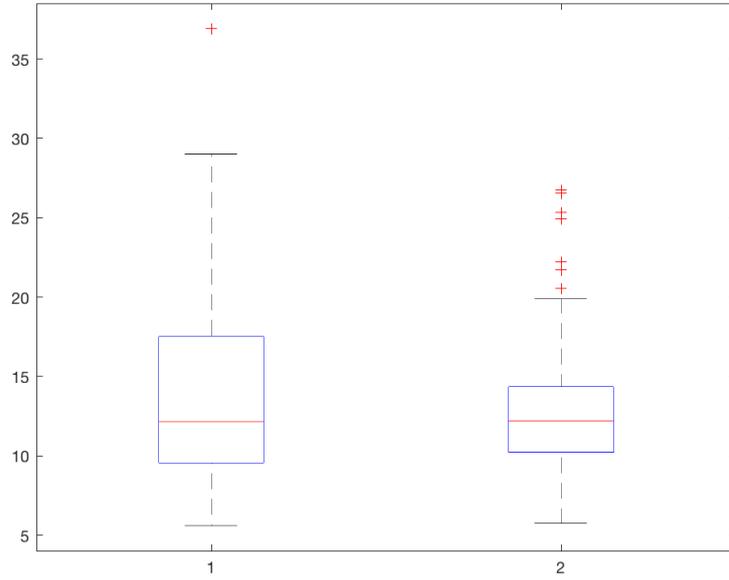
(a) Observed processing times (1) and equilibrium processing times (2)
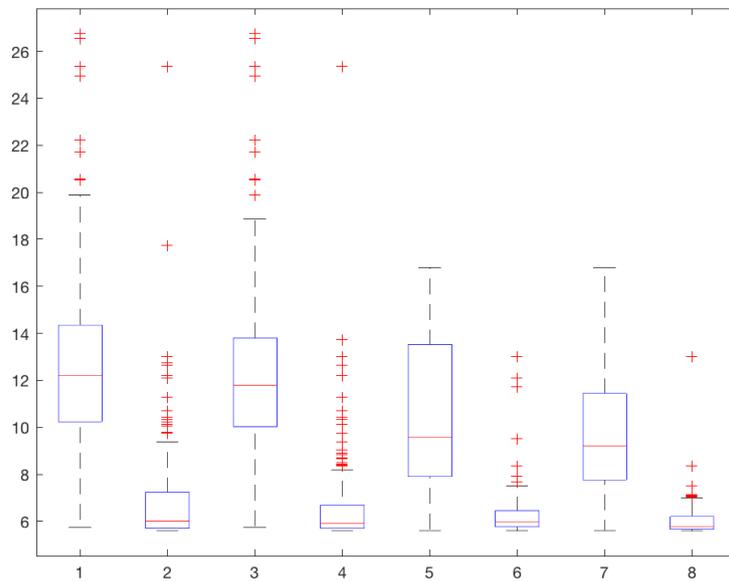
(b) Processing time distribution of courts across the different 8 policy scenarios

Figure 6: Boxplots of the processing time

(a) Scenario 1   (b) Scenario 2

(c) Scenario 3   (d) Scenario 4
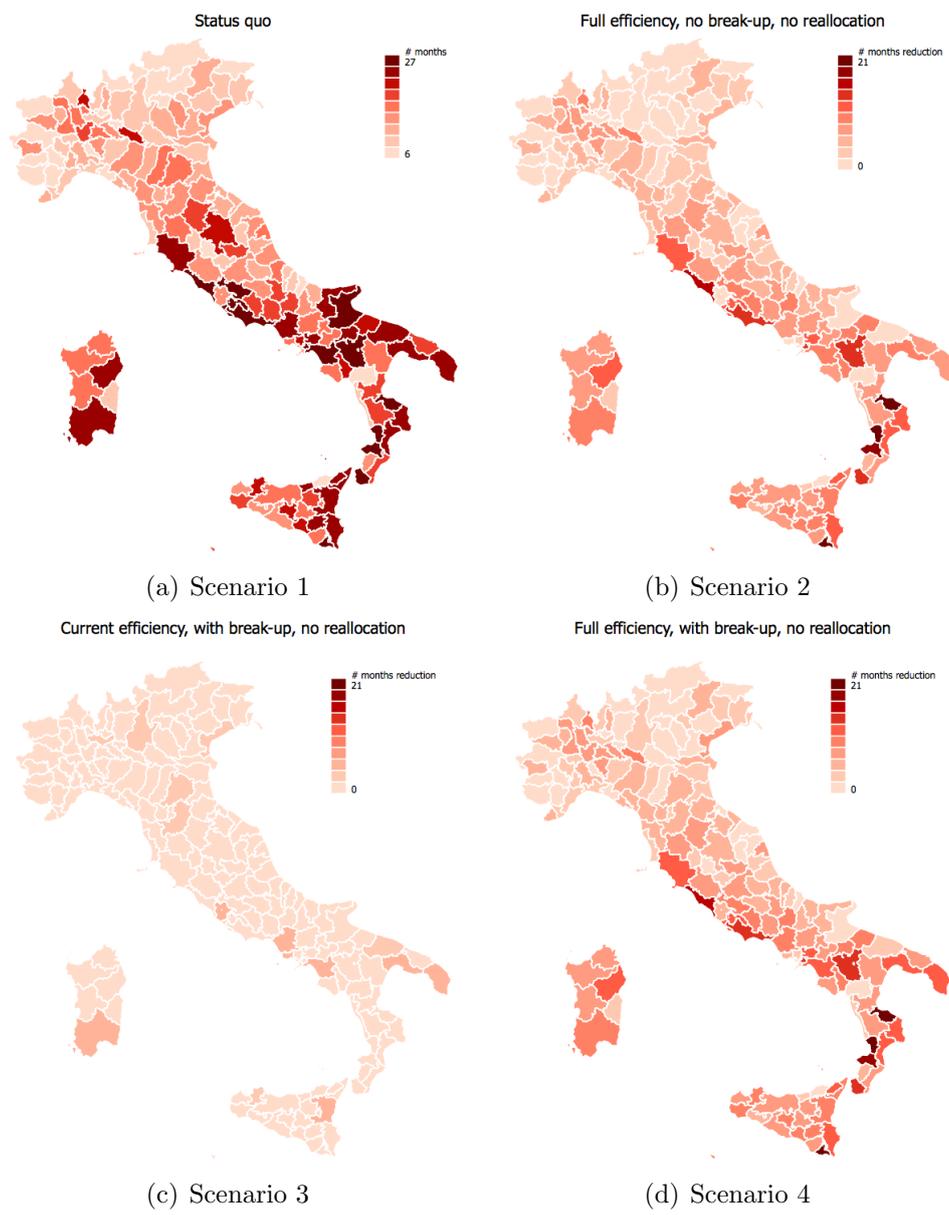
Figure 7: Policy scenarios - Without judge reallocation

Current efficiency, no break-up, with reallocation

Full efficiency, no break-up, with reallocation

(a) Scenario 5

(b) Scenario 6

Current efficiency, with break-up, with reallocation

Full efficiency, with break-up, with reallocation

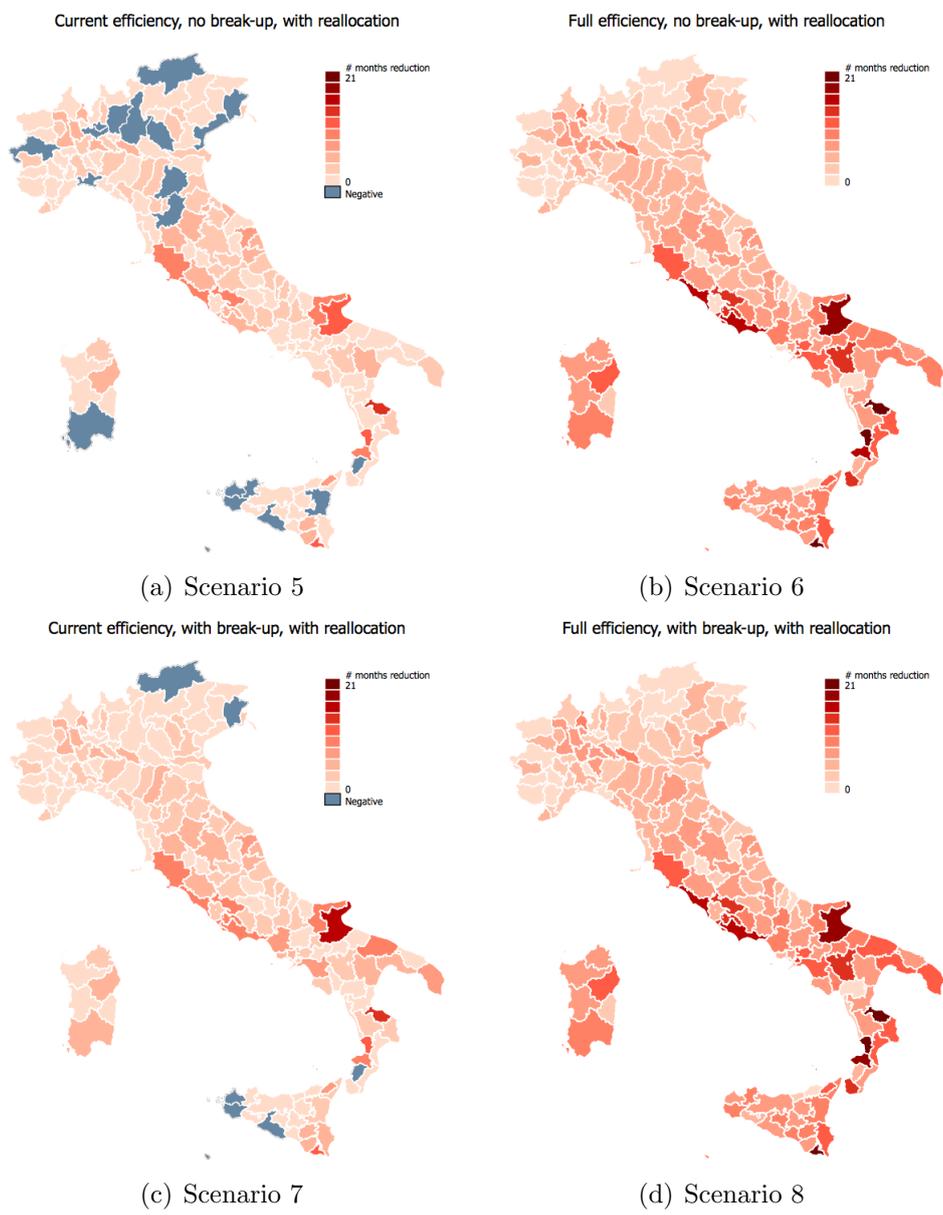(c) Scenario 7

(d) Scenario 8

Figure 8: Policy scenarios - With judge reallocation

Table 4: Inputs and outputs (1,320 obs.)

Variables	2005	2006	2007	2008	2009	2010	2011	2012	2005-12
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
	(St.dev.)	(St.dev.)	(St.dev.)	(St.dev.)	(St.dev.)	(St.dev.)	(St.dev.)	(St.dev.)	(St.dev.)
	min.	min.	min.	min.	min.	min.	min.	min.	min.
	[Max.]	[Max.]	[Max.]	[Max.]	[Max.]	[Max.]	[Max.]	[Max.]	[Max.]
Outputs									
Defined cases	23,072	22,275	23,402	24,159	24,763	24,461	24,041	24,025	23,850
	(33,829)	(32,675)	(33,668)	(34,434)	(34,733)	(34,269)	(33,885)	(32,650)	(33,838)
	1,626	1,567	1,556	2,023	1,538	1,969	1,908	1,806	1,807
	[272,041]	[254,732]	[250,495]	[256,390]	[260,104]	[259,571]	[264,576]	[255,645]	[254,959]
Inputs									
Judges	37	34	34	34	34	34	34	34	34
	(58)	(51)	(51)	(51)	(53)	(53)	(53)	(53)	(52)
	6	6	6	6	6	6	6	6	6
	[469]	[410]	[410]	[410]	[406]	[406]	[406]	[406]	[408]
Pending cases	27,641	27,616	28,047	27,976	28,131	28,206	28,262	28,029	28,040
	(40,772)	(40,051)	(39,988)	(40,326)	(40,053)	(39,847)	(39,092)	(38,199)	(39,740)
	1,666	1,979	1,894	1,712	1,954	1,970	1,847	1,753	1,894
	[284,209]	[275,040]	[274,464]	[275,042]	[282,851]	[281,646]	[271,278]	[262,097]	[276,720]
Others									
Incoming cases	23,048	22,705	23,332	24,314	24,839	24,517	23,808	–	23,919
	(33,168)	(32,538)	(33,611)	(34,298)	(34,983)	(33,675)	(33,178)	(–)	(23,603)
	1,763	1,716	1,530	1,962	1,550	1,710	1,625	–	1,782
	[247,909]	[244,876]	[251,855]	[264,199]	[261,775]	[252,840]	[265,809]	[–]	[255,280]
Population	356,071	358,371	361,329	363,910	365,699	367,433	359,965	361,729	362,784
	(391,277)	(397,742)	(400,020)	(401,768)	(404,330)	(406,983)	(394,693)	(397,813)	(400,897)
	21,351	21,196	21,068	20,932	20,789	20,692	20,630	20,525	20,884
	[2,563,120]	[2,705,603]	[2,718,768]	[2,724,347]	[2,743,796]	[2,761,477]	[2,614,263]	[2,638,842]	[2,711,376]