The Normative Gap: Mechanism Design and Ideal Theories of Justice

Zoë Hitzig*

Abstract

This paper investigates the peculiarities that arise when mechanism design is deployed in contexts in which issues of social, racial and distributive justice are particularly salient. Economists' involvement in redesigning Boston's algorithm for allocating K-12 students to public schools serves as an instructive case study. The paper draws on the distinction between *ideal theory* and *non-ideal theory* in political philosophy and the concept of *performativity* in economic sociology to argue that mechanism can enact elaborate ideal theories of justice. A *normative gap* thus emerges between the goals of the policymakers and the objectives of economic designs. As a result, mechanism design may obstruct stakeholders' avenues for normative criticism of public policies, and serve as a technology of depoliticization.

Keywords: mechanism design, school choice, ideal theory, performativity

*Department of Economics, Harvard University, 1805 Cambridge Street, Cambridge MA 02138. Email: zhitzig@g.harvard.edu. URL: https://scholar.harvard.edu/hitzig

1 Introduction

In a 2003 paper in the *American Economic Review*, economists Tayfun Sönmez and Atila Abdulkadiroğlu framed the thorny problem of assigning K-12 students to public schools in game theoretic terms (Abdulkadiroğlu and Sönmez 2003). In addition to formulating school assignment as a problem from the branch of microeconomic theory known as *mechanism design*, they analyzed existing school choice allocation systems in Boston, Columbus, Minneapolis, and Seattle. The economists demonstrated through proofs and propositions that the existing systems have "serious shortcomings" and that adopting a different mechanism could "provide a practical solution to some of these critical school choice issues" (Abdulkadiroğlu and Sönmez 2003: 742). A reporter from the *Boston Globe* summarized the broad policy implications of the economists' work:

> For more than two decades, policymakers have devoted enormous amounts of attention to various ways to assign students to schools, sparking philosophical debates, charges of racial and economic discrimination, and tangled court battles—all of which have played out with particular drama in Boston. But the authors say their work....is the first rigorous examination of how best to do the matching.
>
> (Cook 2003)

Shortly after their landmark paper appeared, Strategic Planning Manager of Boston Public Schools (BPS), invited Abdulkadiroğlu and Sönmez along with economist Al Roth and doctoral student Parag Pathak to an October 2003 meeting to hear about problems with the current system of allocation (Abdulkadiroğlu et al. 2005b). In 2005, BPS implemented a new system of assigning students to schools based on the economists' recommendations.[1]

What changes did the economic theorists recommend, and how did their proposal relate to the issues that had tormented Boston's public school system for decades? In this paper, I investigate the peculiarities that arise when mechanism design is deployed in contexts in which issues of social, racial and distributive justice are particularly salient. Using the 2005 redesign of the BPS student assignment algorithm as a case study, I argue that mechanism design does not simply enact economic efficiency, but can also enact ideal theories of distributive justice. It is in this enactment that it is possible to identify mechanism design's limitations. When an ideal theory of justice is enacted through mechanism design, a *normative gap* emerges between the normative goals of the policymakers and the normative theory underlying economists' designs. As a result, mechanism design may obstruct stakeholders' avenues for normative criticism of public policies, and serve as a technology of depoliticization.

This paper contributes to an emerging literature commenting on recent advances and applications in theoretical and experimental mechanism design. These advances lie at the intersection of economics, computer science and operations research and are often united under the headings of "design economics" or "market design."[2] A range of historians, philosophers and sociologists of science have brought their critical perspectives to bear on the introduction of mechanism design—supplemented with experiments and computing techniques—into policymaking. For example, a handful of works explore the so-called "performativity thesis" in economic sociology (Callon 1998b) in relation to market designers' activities through studies of the electromagnetic spectrum auctions (Mirowski et al. 2007) and the design of utilities markets (Breslau 2012, 2013). These works suggest that the rules and procedures drawn up in economic theory create a performative frame, formatting actors and institutions in a way that makes them more closely

[1] I refer to the group of economists who advised BPS as Abdulkadiroğlu et al. for simplicity, not because Abdulkadiroğlu had any special contribution. Several of them were already involved in designing the school allocation system in for High Schools in New York City when they were enlisted in Boston (Abdulkadiroğlu et al. 2005a).

[2] Al Roth's 2002 Fisher-Scultz lecture titled "Economist as Engineer: Game Theory, Experimentation and Computation as Tools for Design Economics" arguably solidified the identity of the nascent field.

resemble economists' theoretical constructs. Others have argued that the distinctive combination of theoretical models and experimental techniques in market design demand a novel understanding the limitations and applicability of rational choice and economic theory (Guala 2001; Alexandrova 2008). In addition, several recent papers discuss how market designers respond to settings in which ethical, political and economic issues are intertwined (Roth 2007; Li 2017; Hitzig et al. 2019).

This paper represents a substantial departure from previous literature in two main ways. First, I focus on applications of design economics in which social justice issues are particularly salient. Second, this paper draws a novel connection between ideal theories of justice in political philosophy and the performativity literature in economic sociology. By laying out an analogy between mechanism design and ideal theories of justice, I show that the performance of *economic* theory may also be a performance of *political* theory. The performativity thesis thus has stakes beyond those articulated in previous works.

The essay proceeds as follows. I begin in section 2 with an overview of the historical and political issues at play in the Boston school system, and I introduce a few central concepts from design economics. Then, in section 3, I demonstrate that mechanism design can be understood as instantiating an ideal theory of distributive justice. As such, a *normative gap* emerges between the goals of the designers and the policymakers. Then, section 4 demonstrates the stakes of the normative gap by drawing on criticisms of ideal theory and the concept of performativity. I conclude with a brief discussion of practical implications of the normative gap, and directions for future work.

2 Background

In this section, I discuss the historical and political context of Boston public schooling, which serves as a prologue to the 2005 BPS redesign. I also introduce the concepts from mechanism design required to understand the case at hand.

2.1. Prologue

Legally speaking, segregation in public schools in the United States ended with the landmark Supreme Court ruling *Brown v. Board of Education* in 1954 (Brown v. Board of Education of Topeka 1954). But the years following that case were marked with struggles over integration. Boston is especially infamous for its handling of desegregation in the school system in the decades after *Brown*, and indeed became a symbol of "white backlash" to civil rights era reforms.[3] In 1972, the Boston chapter of the National Association for the Advancement of Colored People (NAACP) filed a class action lawsuit against the Boston School Committee, charging the group with maintaining racial segregation in their schools in violation of the Thirteenth and Fourteenth amendments (Morgan v. Hennigan 1974). Judge Wendell Garrity Jr. eventually ruled, in 1974, that the defendants had indeed contributed to racial segregation in public schools.

After the ruling, which mandated that all schools be racially balanced, Judge Garrity put forward a controversial plan: to desegregate public schools through compulsory busing. Between 1974 and 1988, students from predominantly black neighborhoods were bused to schools in predominantly white neighborhoods and vice versa in order to achieve racial balance in public schools. Throughout the busing period, the policy earned fierce criticism from all sides and was blamed for inciting riots, protests and other violent incidents (Formisano 2004).[4] Judge Garrity's active role in the busing controversy was vehemently criticized. Over a decade after busing

[3]For discussions of the Boston's struggle with desegregation in their school system, find book length treatments in Reid (1974), Sheehan (1984), Tager (2001), Formisano (2004), Naimark (2012), and Delmont (2016).

[4]Some argue that busing contributed to "white flight" from Boston-proper to surrounding suburbs. Over the busing period, the number of students in the Boston school district decreased from around 100,000 students to just 57,000, and the proportion of white students decreased from 65% of total enrollment to just 28% (Formisano 2004: 16).

ended, a report out of the Hoover Institute summarized Garrity's role and subsequent effects on the system, "A federal judge's experiment in social engineering has unraveled neighborhoods and frustrated black achievement" (Richer 1998: 1).

Boston began to petition Garrity to adopt a controlled choice procedure starting in 1984 (Boston School Committee, Office of the Secretary 1984). Under a controlled choice plan parents have explicit choice over the schools their children might attend in or outside their residential neighborhoods. No student is guaranteed a spot at any particular school—rather, with parent preferences in hand, the district officials assign students in a way that compromises among three objectives: (i) to give families some choice over which school their children would attend; (ii) to administer a fair procedure; and (iii) to ensure a balanced distribution of students in terms of various demographic characteristics (Alves and Willie 1987). In 1988, Boston implemented a controlled choice procedure with limited busing. Boston's controlled choice plan was in effect until 1999 when, freed from the necessity to balance school enrollment across demographic characteristics, an *immediate acceptance mechanism* was adopted.[5] Under the new system, the 60,000 students in Boston's Public Schools in Kindergarten through 12^{th} grade submitted a ranked list of no more than five schools in the spring before Kindergarten, 1^{st}, 6^{th} and 9^{th} grades (Abdulkadiroğlu et al. 2005b). Under the immediate acceptance mechanism, BPS assigned students to their first choice school whenever possible.

For over-demanded spots, allocation occurred via a system of priorities. First priority went to students who already had a sibling at that school. Next, students who lived in that school's "walk-zone"—areas which would allow students to walk to school—got priority. Then, the final seats were allocated by random lottery. The first stage ended when either no seats remained or no student remained who listed that school as their first choice. At the second stage of the process, after each school considered all the students who had listed that school as their first choice, all of the schools with remaining spots considered all of the students who listed them as their second choice, and allocated the seats in priority order. The process continued (third stage, fourth stage, etc.) until every submitted choice had been considered. When the process terminated, any students who were not assigned to any of their choices were assigned to the closest school with available spots.[6]

Though the immediate acceptance mechanism did not stir the level of controversy of the busing system, it nonetheless had its flaws. For the unlucky students who did not get their first choice school, their second choice school was likely to already be filled with students who listed it as their first choice. So, there was a chance that a student would fail to get a spot at her second choice school that would have been available to her if she had listed her second choice school as her first choice. And by the time she considered her third choice school, her third choice would already be already filled with students who listed it as their first or second choice...and so on. Thus, there was an incentive for families to "game" the system by writing down a "strategic" ranking of schools that may not reflect their true preferences.

Indeed some parents strategized. Others did not. The result was that parents who did not strategize were less likely to receive their top choices. For example, Abdulkadiroğlu et al. (2006) show that in 2003, 20% of parents listed two over-demanded schools as their top two choices. Ranking two over-demanded schools as top choices is an example of a particularly naive strategy. Of those 20% who filled out the form in this way about 27% ended up unassigned. Further, Abdulkadiroğlu et al. (2006) also show that the 20% of parents who filled out the form in this way tended to come from lower income brackets and less wealthy neighborhoods.

In technical terms, the fact that some can "game" the immediate acceptance mechanism shows that it is not *strategy-proof*. In this context, "gaming the system" has a highly specific

[5] In the school choice literature, this mechanism is more commonly referred to as the *priority mechanism* or sometimes the *Boston mechanism*. I use *immediate acceptance* in order to keep terms and concepts clear. The first discussion of the *priority mechanism* was in Roth (1991), in reference to the process that had been used to match medical graduates to internships in the United Kingdom.

[6] As described in Abdulkadiroğlu et al. (2005b); Pathak and Sönmez (2008).

meaning, which I will dwell on in section 3. Abdulkadiroğlu and Sönmez (2003) note that the immediate acceptance algorithm is not strategy-proof. In talks with BPS, they (together with Roth and Pathak) made an empirical case to BPS that strategizing occurred (Abdulkadiroğlu et al. 2006). When they proposed redesigns of the mechanism for assignment, they considered only strategy-proof mechanisms. Reflecting on their design decision, they write, "As far as we know, it is the first time that 'strategy-proofness'...has been adopted as a public policy concern related to transparency, fairness, and equal access to public facilities" (Abdulkadiroğlu et al. 2006: 2). While other applications of mechanism design had taken strategy-proofness into account as an incentive constraint, in the Boston case, strategy-proofness was the primary goal: "it came to be seen as a criterion of equal access for families with different degrees of sophistication about the system"(Abdulkadiroğlu et al. 2009: 1956).

What exactly did the economists change about the mechanism in Boston? In order to explain the policy change, I present a brief overview of concepts from mechanism design pertinent to the school choice case.

2.2. School choice as a mechanism design problem

The founding paper of school choice—and arguably of market design—is David Gale and Lloyd Shapley's 1962 paper "College Admissions and the Stability of Marriage" (Gale and Shapley 1962). Gale and Shapley present a *two-sided matching model* in which there are two sets of agents—students and schools—and each agent has preferences over agents in the other set. The canonical *school choice problem* as presented in Abdulkadiroğlu and Sönmez (2003) adapts several key mathematical results from the literature initiated by Gale and Shapley to analogues in assigning K-12 students to public schools.[7]

The solution to a school choice problem is a *matching* that assigns each student i to a school, say s. The matching is thus referred to as a function $\mu : I \to S$, where $\mu(i)$ thus refers to the school assignment for student i. There are two concepts that apply to matchings that are relevant for our purposes—*stability* and *efficiency*.[8] A matching is *stable* if no unmatched student-school pair (i, s) exists such that student i prefers school s to her assignment *and* she has higher priority than some other student who is assigned a seat at school s.[9] The concept of *efficiency* as applied to matchings is the familiar definition of Pareto efficiency. A matching is *efficient* when no student could be matched to a more preferred school without causing a change in the allocation such that another student is matched to a less preferred school.

While the concepts of stability and efficiency apply to matchings, *strategy-proofness* is a concept that applies to the matching *mechanism*. A mechanism is a systematic procedure that selects a matching for any school choice problem. In other words, the mechanism takes as input any arbitrary school choice problem—that is, some set of student preferences and a set of school priorities—and constructs a matching. Thus, a mechanism φ is a function from an arbitrary school choice problem to a matching. It takes as input a set of schools and a set of students, along with the priorities and preferences of those schools and students, and outputs an allocation of students to schools. Under a *strategy-proof mechanism*, a student's best assignment will be the one attained by submitting truthful preferences.

Thus, mechanisms can have various desirable properties. They can reliably construct stable

[7] A key departure from Gale and Shapley's original model is that schools are passive rather than active agents. In the school choice problem, schools do not have preferences over students based on student talent, ability or personality. Though schools do not have preferences over students, they do have a system of *priorities* that is imposed by state or local laws. Abdulkadiroğlu and Sönmez (2003) showed that priorities can be treated like preferences in the original two-sided matching model.

[8] There are many other concepts that apply to both matchings and mechanisms that I do not discuss in this paper.

[9] Strictly speaking, the counterpart to stability in the school choice problem is actually what Abdulkadiroğlu and Sonmez call *elimination of justified envy*. They make a distinction because stability applies two-sided problems, while the school choice problem is framed as a one-sided problem.

matchings when such matchings exist, they can reliably construct efficient matchings when such matchings exist, and they can be strategy-proof. Note that since the school choice problem is a two-sided matching problem, the concepts of *stability*, *efficiency*, and *strategy-proofness* can apply to students, schools or both. In the case of Boston's school choice mechanism, schools are not strategic players—priorities are defined by local laws rather than by individual school administrators. So the main concepts of interest are stability, efficiency and strategy-proofness *for students*. Much of the theoretical literature on school choice revolves around existence and impossibility theorems about which mechanisms can and cannot achieve these and other desiderata for one or other or both side(s) of the market.

The most important example of a strategy-proof mechanism for our purposes is the *deferred acceptance algorithm* defined by Gale and Shapley and proven to be strategy-proof in Dubins and Freedman (1981) and Roth (1982a). The deferred acceptance algorithm, applied to the school choice problem, works as follows. At the first stage, each student "proposes" to her first choice school. The schools that have received proposals consider the proposals according to priority order and tentatively hold as many students as possible under their capacity. The remaining students are rejected. Then, at the second stage (and for all following stages k) all of the students rejected at stage 1 (or $k-1$) propose to their second choice school. The key feature of the deferred acceptance algorithm is that it initially assigns students tentatively, so that students with higher priority can be considered over students who ranked that school first in later stages of the assignment. The procedure ends at the stage when no student proposal is rejected. The tentative assignments then become final assignments.

When Abdulkadiroğlu et al. were brought in to help BPS redesign their allocation method, the economists suggested two alternative procedures and settled on the one based on the famous deferred acceptance algorithm.[10]

3 Engineers or philosopher-kings?

How are we to understand BPS's choice to implement a deferred acceptance algorithm in the context of Boston's historically fraught school system?[11] Roth's (2002) self-description suggests that economic theory allows economists to "engineer" institutions for policymakers. He gestures to the relationship between physics and engineering:

> Consider the design of suspension bridges. The simple theoretical model in which the only force is gravity, and beams are perfectly rigid, is elegant and general. But bridge design also concerns metallurgy and soil mechanics, and the sideways forces of water and wind. Many questions concerning these complications can't be answered analytically.
>
> (Roth 2002: 1342)

I suggest that in the Boston case, many of these "complications [that] can't be answered analytically" are principally normative in nature.[12] Thus, I seek to understand how the normative

[10]The other mechanism they suggested was another strategy-proof mechanism called the *top trading cycles* mechanism. The top trading cycles mechanism was originally introduced in Abdulkadiroğlu and Sönmez (2003) and is an extension of a mechanism introduced in Shapley and Scarf (1974) and proved to be dominant strategy incentive compatible in Roth (1982b). I will not explain the top trading cycles mechanism here, nor will I get into the details of why the top trading cycles mechanism was rejected.

[11]Throughout this section, when I refer to BPS's application of mechanism design, I am referring solely to their 2005 implementation of a strategy-proof mechanism. Since 2005 there have been further modifications to the original policy, which I will discuss to some extent in section 4. But for present purposes I focus on the isolated policy change.

[12]There are certainly other sorts of complications in this case. Notably, there are psychological concerns that substantially complicate the market designers' task.

"complications" in this setting are handled by economists who think of themselves as engineers.[13]

After giving a brief summary of standard concerns, I argue that the BPS case highlights two further issues with the engineering analogy. Then I show, through an inference to best explanation, that the economic theory used in the Boston case can be understood as enacting a version of an *equal opportunity for welfare* principle of distributive justice. On this understanding, the economists do not resemble engineers as much as they resemble political philosophers given the opportunity to make a thought experiment real.

3.1. Two problems with the engineering analogy

What is implied by the engineer analogy? Engineers combine scientific theory with practical know-how to offer designs that bring about a certain outcome in the real world—e.g. a structurally sound suspension bridge. Engineering a solution proceeds through the following stylized process. First, engineers translate salient aspects of a practical problem into their theoretical analogues. Then, engineers solve the theoretical problem, and find the best real-world approximation to the theoretical solution by taking into account real-world constraints such as the materials available, budget constraints and the objectives of the design. Engineers' technical solutions to practical problems are thus widely understood as *value-free* in an important sense.[14] When engineers do make value judgments, they do so in order to fill in an *instrumental* gap between theory and practice. Their value judgments are directly related to their technical expertise: they do not require recourse to considerations outside the engineer's domain of expertise, domains such as ethics and politics. Engineers choose means to a clearly defined end—when they make value judgments, they do not play any role in interpreting or defining the ends themselves.

Standard criticisms of the engineering analogy in economic policy draw attention to the instrumental gap. They suggest that the engineering analogy understates the scope of economists' instrumental reasoning. Notably, a broad family of well-worn objections to the involvement of economists in policy-making centers on economists' near-ubiquitous reliance on the concept of *efficiency* to make statements about welfare.[15] Armed with neutral principles such as Pareto efficiency—and related metrics such as Kaldor-Hicks efficiency and Bergson-Samuelson social welfare functions—economists think of themselves as engineers in the realm of policy. They apply their narrow concepts of welfare and efficiency to the evaluation of optimal policy, much the way a physicist applies Newtonian physics to the design of suspension bridges. They provide the means to help policymakers achieve their ends, while remaining silent on the ends themselves.

But, the standard objection goes, policy goals are much more normatively complex and multidimensional than the goal of suspension bridge design, so economists necessarily must engage in some reductive interpretation of social goals to apply their tools. The standard objections take issue with the way in which economists fill in the *instrumental* gap between theory and practice. Understanding how to factor in various normative "complications" that arise in evaluating policies is not, as Roth suggests above, like factoring in "metallurgy and soil mechanics, and the sideways forces of water and wind." Economists must interpret the incomplete description of policymakers' ends in order to come up with a means to those ends using their concepts of preference, welfare, and efficiency. In doing so, the instrumental reduction

[13]Though the economist-as-engineer identity was solidified for market designers in particular with Roth's (2002) manifesto quoted above, the metaphor is not new. Economists have been described in many contexts as "social engineers."

[14]Of course engineering is not entirely value free. Whenever theory and practice combine, a practitioner guided by theory must make *some* value judgments (Schon 1983). Moreover, there is much disagreement about what exactly value-freedom implies, and whether its a useful ideal. I wish to skirt these discussions, but take it as intuitive that this view is widely held.

[15]For an introduction to the family of objections I discuss in this and the following paragraph, see survey texts such as Hausman and McPherson (2006), Hausman (2011) or Reiss (2013).

to the concept of efficiency traffics in the economists' own ideas about what *should* be. Such arguments are too numerous to cite here, but Daniel Hausman and Michael McPherson concisely capture a common thrust: "Economists are not only social engineers contributing to policy in the way that civil engineers contribute to policies concerning dams" (Hausman and McPherson 2006: 306). Rather, they continue, "normative economics attempts to appraise policies, even if usually from a limited point of view, and evaluative thinking is in practice unavoidable in order to formulate well-defined questions for positive inquiry" (306). These issues highlight ways in which economists' domain of expertise is narrower than their diagnoses of and solutions to policy problems suggest.

There are two further features of the school choice case that suggest the engineer metaphor is misleading. These features, taken together, demonstrate the possible emergence of a *normative* gap in market design that is not present in, say, civil engineering. The first feature (F1), which I introduce briefly here and return to in section 4, is that the economic theory is *enacted*. The stakeholders in the school choice redesign are formatted as individuals with well-ordered preferences, and thus "perform" the desirable properties of the mechanism.[16] This enactment is unlike the simplifying assumptions made in engineering, which typically do not re-format *agents* to conform to stylized assumptions.

Features of School Choice as "Engineering"

(F1) Theory is *enacted* by participating agents.

(F2) Theory is grounded in a more *elaborate normative theory* than maximization of efficiency.

The second unusual feature of this case (F2), which I will discuss for the remainder of this section, is that the mechanism design used relies on more complex normative concepts than simple efficiency. Economists do not simply fill in an instrumental gap when they propose policy solutions—instead, they introduce a normative gap. In contrast to the instrumental gap, which arises when economic theory is normatively *limited*, the normative gap emerges because economic theory is normatively *capacious*. The various desirable properties that solutions to mechanism design can have are often referred to as "normative properties" of different mechanisms. This plurality of normative properties indicates that market designers' normative work does not define the best mechanism as the "efficient" mechanism. Choosing between mechanisms with different normative properties requires a normative judgment. It is illustrative that a retrospective paper by two of the economists involved in the redesign of Boston's mechanism was titled "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism" (Pathak and Sönmez 2008). This paper concludes:

> Boston Public Schools... identified a fairness rationale for a strategy-proof system. In this paper, we examined this intuitive notion and showed that the Boston mechanism favors sophisticated parents at Pareto-dominant Nash equilibrium, providing formal support for BPS's position.
>
> (Pathak and Sönmez 2008: 1646)

In what sense does economic theory provide formal support for BPS's fairness rationale? If economic theory can "level the playing field," it must do more than collapse a multidimensional social goal into the language of efficiency. In fact, economic theory seems to map the multidimensional social goals into an egalitarian theory of justice.

[16]There are other ways in which mechanism design, when used to construct institutions in the real world, goes beyond the standard notions of applying theory. For example, in a philosophical discussion of the use of mechanism design to construct bidding procedures for electromagnetic spectrum auctions Alexandrova (2008) argues "the use of game theory for auction design simply does not count as theory application" (Alexandrova 2008: 384).

3.2. Inference to best explanation: Equal opportunity for welfare

I have suggested that the engineer metaphor is misleading in the BPS case because of an unusual feature of mechanism design. There are many desirable normative properties under consideration, and selecting among them can require elaborate normative theories (F2). What distinguishes the BPS case and related uses of mechanism design from other instances of economists' involvement in public policy is that there is a *normative gap* between the language of theory and policy in addition to the usual *instrumental gap.* Economists do not just develop the means to an end but also define that end in their own terms *and* come up with the means to achieve it.

So what exactly is the end in the BPS case? In order to go about understanding the normative theory implicit in the BPS policy change, I make an inference to best explanation. This explanation is one possible interpretation of how the economists and policymakers might have been implicitly thinking about justice. In order to best explain the trade-offs that economists and policymakers made in recommending and implementing a new algorithm, I take the solution adopted to implicitly endorse a view related to Richard Arneson's *equal opportunity for welfare* account of distributive justice (Arneson 1989, 1991, 1999). I do not claim that they were influenced by Arneson's theory directly, nor that they were consciously thinking about their philosophical position. I do suggest, however, that understanding the economic theory as endorsing a particular implicit normative theory helps us understand the peculiar *normative* role that mechanism design introduced in the Boston school choice case, and further, points to more general conclusions about the use of mechanism design in justice-oriented institutions.

Game theoretic thinking has played a starring role in theorizing about justice in the contractarian tradition.[17] It is useful to contextualize the activity of mechanism design within this tradition. In the context of constructing justice-oriented institutions, mechanism design begins with a distributive theory and then works backward to find the institutions which, given some theory of rational agency, yields the desired outcome. Before beginning to examine what kinds of institutions yield the desired outcome, theorists shape their agents into a particular kind of rational agent with certain kinds of knowledge and information and not others. On the face of it, it may seem that since game theory is so readily employed in theorizing about justice, then mechanism design is well-suited to the *practice* of justice. Mechanism design begins with a distributive principle and locates the institutions, which, under rational choice and a given informational environment will enact it.

With this intellectual history in mind, the Boston school choice case begins to resemble a thought experiment from political theory. On my inference to best explanation, some distributive principle was chosen, and then a mechanism was designed to enact that principle given some assumptions about the rationality of participants. I will return to a broader comparison between political philosophy and market design in section 4.1. For now I focus on the distributive principle chosen. As Abdulkadiroğlu et al. report in their 2005 paper, after the economists had proposed several possible mechanisms with different normative outcomes, BPS spent some time "thinking through their philosophical position on the trade-off between stability and efficiency" (Abdulkadiroğlu et al. 2005b: 371). Eventually, they chose a version of the deferred acceptance algorithm for its normative properties of stability and strategy-proofness.

What exactly is the distributive principle settled on by BPS, which led to their choice? While general theories of justice aim to characterize principles of justice to guide the design of *all* social institutions, the case at hand requires a local version of such distributive principles.

[17] Social contract theorists have made extensive use of axiomatic bargaining theory to show how rational agents come to agree on a particular distribution of the benefits and burdens of social cooperation (Braithwaite 1955; Harsanyi 1955; Rawls 1958, 1971; Gauthier 1986: among others). Often these theorists of justice have drawn on contemporary game-theoretic literature. For example, Harsanyi, Gauthier and Rawls employed bargaining solutions proposed in Nash (1950), Kalai and Smorodinsky (1975) and Kalai (1977), respectively, to support their chosen distributive principles. See also Barry (1989) and Binmore (1994, 1998).

Theories of justice aim to answer the question

> *How can a group organize the basic structure of society according to principles that could be justified to all?*

and introduce various sorts of theoretical constructs to do so. In the BPS case, the distributive question is a local version applied to a single institution:

> *How can BPS assign students to public schools according to principles that could be justified to all?*

The central planner in the BPS case, namely the school committee, is in *direct control* of only one distributive outcome. Of course that direct control feeds into *indirect control* because education is an input to many other outcomes that obtain over the course of one's life. While these ramifications are considered in choosing a distributive principle for the school system, the distributive principle itself only applies to outcomes over which the school committee has direct control, i.e. the allocation of students to schools.

Given this narrowed scope of social justice, I argue through an inference to best explanation that the distributive principle implicit in BPS's policy change is closest to *equal opportunity for welfare*, as advanced by Arneson. Arneson writes, "people share equal opportunity for welfare just in case there is some time at which their opportunities are equal and if any inequalities in their opportunities at later times are due to their voluntary choice or differentially negligent behavior for which they are rightly deemed personally responsible" (Arneson 1989: 86). The theory relies on a definition of welfare as preference-satisfaction and of ideally considered preferences.[18] I now demonstrate how BPS enacts a variant of equal opportunity for welfare by walking through the procedure for assigning students to schools in normative terms using a minimal example.[19]

Consider a school choice problem in which there exist only three students i_1, i_2 and i_3 and three schools s_1, s_2 and s_3. The set of preferences P of students over schools is as follows.

Student	1st	2nd	3rd
$i_1:$	s_2	s_1	s_3
$i_2:$	s_1	s_2	s_3
$i_3:$	s_1	s_2	s_3

The set of priorities for schools over students are defined by law. As described in section 2, in 2005 these priorities took into account "walk-zones" and sibling attendance.

School	1st	2nd	3rd
$s_1:$	i_1	i_3	i_2
$s_2:$	i_2	i_1	i_3
$s_3:$	i_3	i_1	i_2

BPS elicits all preferences of students over schools P and computes all school priorities of schools over students, π. Then the centralized mechanism φ creates a matching μ between students and schools. The one that BPS chose, the deferred acceptance algorithm, results in the matching that I shall refer to as μ_{DA}. For comparison, consider the matching that results from the immediate acceptance algorithm (used in Boston from 1999 to 2004) which I call μ_{IA}.

$$\mu_{DA} = \begin{bmatrix} i_1 & \text{goes to} & s_1 \\ i_2 & \text{goes to} & s_2 \\ i_3 & \text{goes to} & s_3 \end{bmatrix} \quad \mu_{IA} = \begin{bmatrix} i_1 & \text{goes to} & s_2 \\ i_2 & \text{goes to} & s_3 \\ i_3 & \text{goes to} & s_1 \end{bmatrix}$$

[18] In this setting, when I refer to welfare, I refer only to the notion of welfare as preference satisfaction, which is distinct from "welfarism." Welfarism is usually committed to principles of Pareto efficiency—for a forceful defense of welfarism see Kaplow and Shavell (2009).

[19] The minimal example presented here is borrowed from Pathak (2011).

I now illustrate how different mechanisms can have different desirable properties. Notice that μ_{IA} is not stable, because i_2 would prefer to be matched with s_2 *and* i_2 has higher priority at s_2.[20] The immediate acceptance algorithm is also not strategy-proof because if student i_2 had submitted a ranking that put s_2 as her top choice, she would have received s_2 instead of s_3, which she preferred less. Though it is neither stable nor strategy-proof, μ_{IA} *is* Pareto efficient because no individual could be made better off without someone else becoming worse off. Conversely, the deferred acceptance algorithm, which is strategy-proof, produces μ_{DA} which is a stable but not Pareto efficient matching. The matching μ_{DA} is not Pareto efficient because if i_1 and i_2 swapped schools, they would both be better off. However, it is still stable because there is no school-student pair that would rather be matched to each other than to their assignment under μ_{DA}. The *top trading cycles* mechanism, also considered by BPS and mentioned briefly in section 2.3 is strategy-proof and efficient, but not stable.[21]

Mechanism	Efficiency	Stability	Strategy-proofness
Immediate acceptance	✓	X	X
Deferred acceptance	X	✓	✓
Top trading cycles	✓	X	✓

Table 1: Three mechanisms and their normative properties.

The three mechanisms considered by BPS and discussed here—deferred acceptance, immediate acceptance, and top trading cycles—each have *some* desirable properties that mechanisms can exhibit (see Table 1 for summary). Thus BPS must make a normative trade-off when choosing which mechanism to implement. When BPS considered its "philosophical position" concerning various trade-offs between different mechanisms, we can understand them as considering the desirable properties of each mechanism and their philosophical analogues.

Now, I demonstrate how the implicit normative political theory underlying BPS's choice of the deferred acceptance algorithm over the others can be best explained as a version of Arneson's equal opportunity for welfare principle of distributive justice. Arneson distinguishes two dimensions along which interpretations of the egalitarian ideal differ. Along one axis, interpretations are after *straight equality* or *equality of opportunity*. On the other axis are bases for measuring distributive shares, with *resources* on one end and *welfare* at the other (Arneson 1989). Equal opportunity for welfare, then can be understood in contrast to other possible combinations. An opportunity is defined as a prospect for acquiring something if one seeks it.[22] In the school choice example, an opportunity is a chance to attend a certain school. In order for equal opportunity for welfare to obtain among students seeking positions at schools, every student must have equal opportunity to satisfy her preferences over schools she would like to attend.

Arneson's view also takes into account the ability of various people to "negotiate" the options presented to them, which in the BPS case translates into a concern about "strategizing." Arneson writes that people might face the same array of options and "yet differ in their awareness of these options [and] their ability to choose reasonably among them" (Arneson 1989: 85). He argues that people must face an *effectively* equivalent array of options, which could obtain in three different ways. First, options can be effectively equivalent if the options are equivalent and the individuals are identical in their ability to negotiate options. Second, if options are not equivalent but they exactly counteract inequalities in ability to negotiate, then options are

[20] This is a slight abuse of terminology—stability and efficiency are concepts that apply to *matchings*. When I refer to a *mechanism* as stable or efficient I mean that the mechanism guarantees a stable or efficient outcome. This terminological elision is standard in the economic literature.

[21] Kesten (2010) demonstrates that there is no strategy-proof mechanism that creates both efficient and stable outcomes whenever they exist.

[22] As Arneson writes, "each must face an array of options that is equivalent to every other person's in terms of the prospects for preference satisfaction it offers" (Arneson 1989: 85).

effectively equivalent. A third way in which options can be effectively equal is for the options to be equivalent and for inequalities in negotiating abilities to be the result of decisions for which individuals are personally responsible. In the case of school choice, the fact that some played strategically and some do not under the immediate acceptance algorithm was seen as unfair. Under Arneson's framework, the different abilities to strategize suggested that none of the three possible ways for options to be *effectively* equivalent obtained. Thus, BPS chose an algorithm under which all individuals have the same options and the same ability to negotiate the options. In other words, the need to satisfy one of the ways of creating *effectively* equivalent options can be understood as BPS's rationale for adopting a strategy-proof mechanism.

Thus I have demonstrated that the school choice algorithm selected by BPS was selected on normative grounds that extend far beyond economists' usual common denominator of "efficiency." Through an inference to best explanation, I suggested that the choice of the deferred acceptance algorithm over other possibilities can be understood as instantiating an implicit normative political theory about what constitutes a just allocation of students to public schools. On this view, a just allocation process provides each participant an equal opportunity for welfare, where welfare is defined as preference-satisfaction.

I do not mean to suggest that equal opportunity for welfare à la Arneson is the only possible normative theory underlying the decision to adopt the deferred acceptance procedure. I simply offer an inference to best explanation, focusing on the prominence of *strategy-proofness* as a normative criterion.[23] By focusing on the *stability* of the chosen algorithm, one could offer a different underlying normative theory. I offer a sketch of an alternative here. Stability is closely related to envy-freeness, which is a distributive ideal in many theories of justice or fairness (Varian 1975; Dworkin 1981a,b: notably). Stability, in Pathak's words, "embodies a notion of fairness: a student should not envy another school over her assignment and have a higher claim to that school" (Pathak 2011). Since the mechanism ensures a stable outcome—an outcome in which there is no "justified" envy, i.e. no student would prefer to be matched to a different school at which they also had higher priority—the allocation of students is envy-free.[24] For the purposes of my argument, it is more important to see that *some underlying normative theory* exists than to precisely pinpoint the content of the normative theory.

4 The normative gap

I now turn a critical eye on the use of mechanism design in constructing BPS's student allocation algorithm. In section 3, I introduced two unusual features of the BPS redesign: that the economic theory is enacted in the school system (F1) and that it draws on an elaborate, but unarticulated, normative framework (F2). Putting (F1) and (F2) together suggests that mechanism design *enacts a theory of justice*. There is a normative gap between the implicit normative theory of the mechanism and the actual demands of justice. In this section I further describe the normative gap, and discuss what is at stake when it emerges. In order to do so, I draw on the distinction in political philosophy between ideal and non-ideal theories of justice, and on the concept of performativity from economic sociology.

4.1. Ideal theory, enacted

What role can or should ideals play in arranging actual institutions? This question is at the crux of the disagreements in political philosophy around *ideal theory* and *non-ideal theory*. While the term ideal theory has been used in a number of ways, I use the original definition in *A Theory*

[23]Recall Pathak's statement that the 2005 redesign represents "the first time that 'strategy-proofness'...has been adopted as a public policy concern related to transparency, fairness, and equal access to public facilities."

[24]The use of random numbers to generate priorities at over-demanded schools complicates envy-freeness somewhat in this context. It may be that the mechanism fails an "everyday" sense of the envy-test, but in its formal construal, it passes.

of Justice (Rawls 1971). Rawls takes ideal theory to be an account of the principles that ought to guide the design of the basic structure of society. Ideal theory operates according to two assumptions absent in non-ideal theory (Rawls 1971, 2001): strict compliance and favorable circumstances.[25] Non-ideal theory, on the other hand, aims to formulate principles that take into account obstacles due to non-compliance and due to unfavorable circumstances. Rawls writes that non-ideal theory "studies the principles that govern how we are to deal with injustice" (Rawls 1971: 8). While critics of ideal theory form a motley group, many dispute Rawls' particular formulation of the proper relationship between ideal theory and non-ideal theory (Mills 1997; Sen 2011; Anderson 2010; Simmons 2010: for example).

What does the disagreement about ideal theory and non-ideal theory have to do with the BPS case? Following my inference to best explanation of the principles underlying the Boston school case, I argue that the mechanism design *enacts* an ideal theory even though it takes place in the non-ideal world. First, the mechanism makes the two idealizing assumptions made in ideal theorizing. It assumes strict compliance—it assumes that students will actually participate in and comply with the allocation procedure. It also assumes favorable circumstances—that students will not be prevented for social, economic or otherwise historically contingent reasons from obtaining information about schools and from rank-ordering the top five schools across the entire district that they would like to attend. Then, with these assumptions in hand, there is an abstract question: How do we allocate students in a way that can be justified to all? In order to answer that question in this realm of abstraction, just as Rawls introduces the veil of ignorance and the original position, mechanism design introduces a formal construct.

In the BPS case, which aims to answer a much more local question about justice, the construct is a strategy-proof mechanism which "'levels the playing field' by diminishing the harm done to parents who do not strategize or do not strategize well" (Boston Public Schools 2005). The strategy-proofness of the mechanism ensures that participants list their "true preferences" rather than trying to "game the system." The strategy-proof requirement thus constrains the way in which participants interact. Participants in the Boston school choice mechanism, as under the veil of ignorance, are *ab initio*—all that matters for acquiring the just distribution is their "true preferences" which are highly idealized.

One might have expected that the mechanism design used in the BPS case, insofar as it is grounded in a distributive principle like equal opportunity for welfare, would belong to the realm of non-ideal theory. After all, non-ideal theory concerns itself with finding distributive principles that handle injustice in the actual world. Boston's public school allocation system had been repeatedly charged with allegations of injustice as discussed in section 2. But, the underlying theory is highly idealized and abstract. The economic theory brought in to "solve" Boston's problems thus enacts a hypothetical construct like those drawn up by theorists of justice like Rawls (and others such as Gauthier and Harsanyi). The mechanism, understood in this way, does not select the just allocation of students to schools in non-ideal circumstances. Rather, it is an artifact of ideal theory. I now turn to a discussion of criticisms of ideal theory and how they come to bear on this enactment.

4.2. Criticisms from non-ideal theory

Do criticisms of ideal theory have bite in this real-world enactment of ideal theory? Or does the enactment, by taking place in the non-ideal world, thus dodge the criticisms of ideal theory? In order to answer this question, I turn to two prominent objections to ideal theorizing from Amartya Sen and Charles Mills, and investigate how each comes to bear on the enactment

[25] That is, it assumes that "everyone strictly complies with" the principles of justice (Rawls 1971: 13) and that "historical, economic and social conditions" (ibid., 47) do not interfere with the possibility of achieving a just society

of ideal theory at hand.[26] I take these criticisms from non-ideal theory to demonstrate the existence of a gap between the ideal of justice enacted by mechanism design and the actual demands of justice. In this section I use "actual demands of justice" as a placeholder, and describe in more depth what these actual demands of justice entail in the next section.

Sen's critique of Rawls centers on a distinction between the *comparative* nature of non-ideal theory and the *transcendental* approach of ideal theory (Sen 2011). Non-ideal theory is comparative in that it allows for comparison of the relative justice of various alternatives. His argument raises issues with two aspects of transcendental theories. First, he argues that the indeterminacy of impartial bargaining devices poses a real problem for locating transcendental theories. The plurality of bargaining solutions indicates the impossibility of finding a "transcendental" theory. Second, and more practically, he argues that abstract ideals of justice are neither necessary nor sufficient for judging the justice or injustice of actual institutions (Sen 2011: 102).[27]

Both of Sen's criticisms of ideal theory relate to the BPS case in insightful ways. Consider the first criticism, and recall that any given mechanism design solution to a school choice problem might have a range of desirable properties. In the BPS case, economists and policymakers specifically focused on the trade-offs between stability, efficiency and strategy-proofness. Sen's criticism in this context calls attention to the fact that other combinations of these desirable features might also serve as distributive principles which, under suitably idealized circumstances, could be justified to all.

Sen's second criticism also instructively applies to the BPS case. It suggests that having a distributive principle for school allocation that can be justified to all in the realm of ideal theory is *neither necessary nor sufficient* for ensuring justice. Sen argues that the transcendental approach and the comparative approach give fundamentally different answers to the question of what justice demands (Sen 2006: 218). The comparative approach looks for "ways and means of advancing justice...in the world by remedying inequities" while the transcendental approach looks for "the simultaneous fulfillment of the entire cluster of perfectly just societal arrangements" (Sen 2006: 218). The deferred acceptance algorithm enacts an arbitrarily chosen transcendental theory. Sen's criticism, applied to our case, suggests it would be misguided to assume that this enactment advances justice.

At times Sen suggests that the pursuit of ideal theory may even be counter-productive—far from playing a necessary role in addressing urgent questions of justice, it may serve to distract from those very questions.[28] A stronger version of this argument has been put forward by Mills with an eye toward racial injustice. Mills has argued that ideal theory obscures the history of racial domination, and thus prescriptions based on contractarian frameworks are necessarily flawed in certain societies (Mills 1997, 2003, 2005). He writes that ideal theory "is really an ideology, a distortional complex of ideas, values, norms, and beliefs that reflect the nonrepresentative interests and experiences of a small minority of the national population" (Mills 2005: 172). Further, he claims that ideal theory "was constructed to evade these [racial] problems" (Pateman and Mills 2007: 107). In order to draw up principles of justice, one must attend carefully to the realities of racial domination, as in non-ideal theory.

Mills' criticisms are useful for understanding the unusual case at hand. In the BPS case, ideal theory is acted out in the real world. As such, it displaces non-ideal theorizing. It abstracts away from racial (and other demographic) realities in order to institute a mechanism that is justifiable to all on its own terms, in a very limited sense. While Mills' assertions about how ideal theory represents "ideology," and how ideal theory was "constructed to evade" racial realities may

[26]While the criticisms of ideal theory are numerous, I choose these two as they are both representative of broader debates and particularly relevant to the case at hand.

[27]The first objection can be understood as an objection to the *existence* of transcendental principles, the second objection is one that supposes their existence and still objects to their *prominence* in evaluating the fairness of existing social institutions.

[28]I refer the reader especially to the entirety of the passage quoted above on p. 218 of Sen (2006), as well as the concluding section of Sen (2011).

strike some as extreme,[29] his provocative criticism provides at the very least a grave warning about the possible dangers of ideal theorizing in the absence of non-ideal theorizing, specifically in the context of racial injustice. Even if ideal theory was not "constructed to evade" racial realities in the causal sense that Mills claims, when ideal theory is enacted as in the BPS case, it might facilitate the evasion of a proper discussion of the real injustice particular to a given situation.

4.3. Performativity and technologies of depoliticization

What exactly are the "actual demands of justice" to which the school allocation system must respond? How does mechanism design wedge a "normative gap" between these demands and the normative properties of the proposed algorithm? In order to articulate how the normative gap emerges, I draw on the performativity thesis from economic sociology.

The concept of performativity, which has its roots in speech act theory (Austin 1962), enthnomethodology (Garfinkel 1956) and actor-network theory,[30] has taken on a variety of meanings when applied to scientific expertise. I focus on the concept of performativity that has been used to describe economists in general,[31] and market designers in particular (Mirowski et al. 2007; Breslau 2012, 2013). To call economics performative is to suggest that economic knowledge does not describe existing phenomena, but rather consists of a set of tools that constructs actors and institutions. In other words, economics provides a performative framing that molds reality in its own image—reality thereby comes to resemble economists' abstractions.[32]

In the BPS case, mechanism design provides a performative framing for the problem of assigning students to public schools. This framing moves the allocation of seats in schools into the realm of ideal theory. Participants are shaped into a certain kind of agent. "Strategy-proofness" becomes the central policy concern related to "transparency, fairness, and equal access to public facilities." The performativity of the economic theory has a blinding effect, as if participants in the mechanism are gradually made into rational agents behind a veil of ignorance, whose knowledge of their actual position in the broader structure of social life is erased.

The framing of households in school choice as agents with well-ordered preferences gives rise to considerations that must be dealt with in the realm of non-ideal theory. These considerations are "the actual demands of justice" relevant to the situation at hand. The public schools in Boston are varied in their resources and opportunities for students, and students at different schools have vastly different outcomes. The choice of the deferred acceptance algorithm changes the overall distribution of students to schools. Are better-off students still landing in the best-resourced public schools? Are there racial disparities that result? These considerations have to do with the collective demands of a school system. The deferred acceptance algorithm serves particular notions of "transparency, fairness and equal access to public facilities"—but these notions can only be achieved *within* the performative framing. Deferred acceptance is "transparent" only to the extent that all participants understand the process. It is "fair" to the extent that prevents a particular form of strategizing. It ensures "equal access to public facilities" if access is understood as opportunity for welfare, i.e. the satisfaction of stated preferences. The performative framing shuts down *other* notions of transparency, fairness and equal access to public facilities, which may arise out of attentiveness to local particulars, the historical injustice of the school system, and the privileging of collective demands for equity.

[29] See Shelby (2003, 2013) for prominent criticisms of Mills' views.

[30] See Latour (2005) for an overview of key concepts of actor-network theory.

[31] See e.g. Callon (1998a,b) for foundational contributions and the edited volumes MacKenzie et al. (2007), Pinch and Swedberg (2008) and Boldyrev and Svetlova (2016) for wide-ranging investigations.

[32] On the face of it, calling market designers "performative" seems trivial. It is rather like remarking that weightlifters have very large muscles. But, the power of performativity as a lens through which to understand market designers' activities lies in precisely identifying how and at what cost economists' definitions of value-laden terms get transposed onto the realities they describe.

The performative framing thus entraps participants in a thought experiment that exists only *ab initio*. Inside this thought experiment, it is difficult for households to assess their actual interests, preferences, and welfare, and how they are or are not represented in the policy change. Participants' concerns about historical injustice, local inequities, and collective demands do not disappear in the BPS case, rather, they are displaced by the performative framing of mechanism design.[33] The normative gap emerges between the framing of student assignment in the language of mechanism design, akin to an ideal theory of justice, and the politics of the allocation of students to public schools. To genuinely address the political stakes requires recourse to non-ideal theory which in turn requires democratic input. The normative gap is hard to close through democratic deliberation because the stakeholders themselves lack the technical understanding that could reframe the interaction.

Why does the so-called normative gap matter? I have not suggested that the economists and policymakers involved in redesigning Boston's school allocation systems were making explicit claims about the suitability of their ideal theory. I have merely provided an inference to best explanation of how their work might be explained as an enactment of ideal theory. So, one might object that my criticism cannot tell us much about the role of economic theory in redesigning the allocation system. In the spirit of the criticisms from non-ideal theory—why not just pay attention to what actually happened in Boston and assess whether the use of mechanism design improved the fairness of the public school allocation system? A hypothetical interlocutor might gesture to two facts to indicate the success of the market designers' work, and thus the irrelevance of the criticism from non-ideal theory. First, recall the civil unrest and the racially-motivated violence that plagued the school system in the latter half of the 20th century. Though BPS still face criticism over the system of allocation, people are not so incensed as to take to the streets. Second, the City of Boston has confirmed the economists' success by continuing to enlist them in refining their system of school allocation in the years since the 2005 redesign.[34]

The mechanism design "succeeds" in these limited ways not because it meaningfully addresses the inequities relevant to the just allocation of resources but rather because it appears to furnish a set of impartial rules and procedures. Mechanism design thus serves as a technology of depoliticization. Because an ideal theory is enacted by the mechanism, opportunities for normative criticism of the rules and procedures are closed off to the stakeholders.

In the wake of desegregation, a select group of experts undermined the public's trust in the school system. That select group of experts might have involved at various points the Boston School Committee, together with Judge Garrity and various consultants. Decision making procedures are especially appealing when expert judgment has undermined public trust. As Theodore Porter argues in his landmark book *Trust in Numbers*, "a decision made by the numbers (or by explicit rules of some other sort) has at least the appearance of being fair and impersonal" (Porter 1995: 8). Recall some of the charges against Judge Garrity that came out in retrospect: "[he] had a paternalistic mentality that all goodness and all knowledge flows from the federal court"(Richer 1998: 5). The criticism that, through busing, a federal judge had taken up "an experiment in social engineering" is indicative of the undermining of public trust.

Gradually, after Garrity, the allocation process became more mechanical and rule-based. The Boston School Committee had historically faced a political, social and economic quandary assigning students to public schools, and their expert judgments had undermined the public's trust in their ability to make decisions. The "controlled choice" plan initiated in 1988 allowed parents to begin to place their "trust" in a mechanism rather than a group of experts. That mechanism became even more rule-based in 1999 when race was dropped as a distributive category. Then, when economic theorists turned their attention to BPS in 2005, they offered

[33]This criticism echoes Breslau's discussion of the stakes of the performativity of economics in the design of electricity markets: "the political conflicts... have not vanished, but they are no longer mediated by the regulatory process. Instead they are now a politics of market design, mediated by an abstract and technical object" (Breslau 2012: 401).

[34]For an archive of past and ongoing policy dialogues since 2003, visit www.bostonpublicschools.org.

an entire mechanical language for *evaluating* the virtues of different allocation algorithms.

The mere fact that school choice had been cast as a problem in mechanism design by Abdulkadiroğlu and Sönmez (2003) made it an attractive alternative to other ways of deciding how to allocate students to schools, simply by virtue of being highly mechanical. The school choice procedures appeared to become more objective in two ways. First, procedures replaced the expert judgments of the school committee—when the "controlled choice" mechanism was put into place in 1988, a procedure replaced individual decision-makers. Second, mechanism design offered a consistent language for *evaluating* the increasingly algorithmic procedures. In 2003, what Abdulkadiroğlu and Sönmez's paper offered was a language for evaluating allocation procedures that offered a veneer of objectivity. This veneer of objectivity makes it easier to justify the procedure to a vested public. Some policy might be chosen not because it actually addresses the issues best but because it appears to offer accountability both in the procedure itself and the evaluation of those procedures.

Consequently, the gap between the normative theory implicit in BPS's use of mechanism design and the justice-based concerns of stakeholders can be understood partially as the result of a bi-level drive toward more seemingly objective rules and procedures. The ideal theory implicit in the mechanism design supplants local ideas about what constitutes justice.[35] Additionally, it is not the normatively best solution—any objective-seeming method would have accomplished the task.

Sen's two part criticism of ideal theory also helps us here. First, recall that there are many ways of reaching an "impartial ideal"—by imposing different constraints on rational agents. The same is true in the school choice case. There are many candidates for a distributive ideal that can be performatively framed through concepts like "efficiency," "stability" and "strategy-proofness." So in the face of this indeterminacy *one must choose* a particular distributive ideal. But the process of choosing an ideal cannot be impartial or objective—it is not aiming for some transcendental truth but rather a fair solution to an actual problem. One would hope, in fact, that the process of choosing among ideals would be highly democratic or collectively deliberative, and therefore closely tailored to the needs of Boston's public. However, the technical evaluation of procedures can easily come to take the place of such democratic deliberation.

One might *still* respond to the critical framing above by suggesting that the current *algorithms* designed by a small group of experts are clearly an improvement over leaving it up to the *judgments* handed down by a few experts. I have not argued otherwise. However, I have argued that the veneer of objectivity is worrisome in conjunction with mechanized procedures that "perform" reality. Mechanism design, used in this way, is a technology of depoliticization, which may obstruct normative criticism in a manner similar to how Mills sees ideal theory as evading unjust realities. Without proper attention to the gap between the normative theory enacted by the mechanism and the actual demands of justice, the gap will persist.

4.4. Epilogue

To further demonstrate how the critical perspectives presented above come together to illustrate the emergence and persistence of the normative gap, I present an epilogue to the story presented in section 2.

I have already called attention to how the priority system—defined by local laws before the mechanism is enacted (subsection 3.2.)—must be decided with the utmost care if the mechanism is to genuinely give students in all neighborhoods an equal opportunity for welfare. When BPS considered different proposals for changing the mechanism in 2003, they were also presented with six different ways of drawing walk-zones, suggested through a community engagement process

[35]This argument further echoes Porter, who writes that mechanical objectivity "aims to supplant local cultures with systematic and rational methods" and in doing so "its resonances are egalitarian" (Porter 1995: 77). In the BPS case, an impersonal, abstract way of thinking about justice has supplanted local ideas about justice. It is not just that its "*resonances* are egalitarian" (emphasis mine) but that its actual content is egalitarian, as well.

(Landsmark and Dajer 2004; Abdulkadiroğlu et al. 2005b; Dur et al. 2018). The redrawing of walk zones would change the demographic distribution of priorities at different schools. However, the only change made by BPS was to change the algorithm for assignment. An attentiveness to collective demands of the school system relies on how the neighborhoods are drawn, but since the mechanism offered a straightforward way of discussing a different kind of fairness, the easiest change to make was a mechanical one—to change the algorithm. The elaborate normative language of the mechanism displaced other sorts of normative questions. The performative framing obstructed deeper questioning of how well the mechanism met collective demands. In the ensuing years, stakeholders began to express concern that the system of walk-zone priority continued to provide an unfair advantage to those who lived in wealthier neighborhoods (Dur et al. 2018). Stakeholders attempted to renew discussions about redrawing neighborhood lines in 2006, the year after the new algorithm was instated, and again in 2009. But the system remained largely unchanged[36]—the normative gap persisted.

In 2012, Boston Mayor Thomas Menino heeded the concerns about walk-zones and pledged a major overhaul of the BPS allocation system. While the concerns about walk-zones were addressed, the normative gap persisted because the economists involved this time around simply offered up a new mechanism—a new performative framing corresponding to a new ideal theory. This mechanism solved a new set of concerns about walk-zones, while the 2005 mechanism change responded principally to concerns about "gaming the system." The mayor commissioned Pathak, one of the original consultants in 2003 who was by 2012 a professor of economics at MIT, to evaluate proposals for overhauling the system. One of Pathak's doctoral students in operations research, Peng Shi, played a key role in evaluating alternative mechanisms for the advisory committee, using computer simulations to analyze the impact of various proposals (Seelye 2013b).

Shi eventually put forward a proposal of his own which scrapped walk-zone priority altogether and would also do away with the final remnants of busing students to faraway neighborhoods as initiated by Garrity in 1974 (Seelye 2013b). Shi writes of his proposal: "I formulated the reform as an optimization problem of finding school-choice menus and priorities that induce the best combination of equity of access, proximity to home, predictability, and community cohesion" (Shi 2015: 1). Under his proposal an algorithm produces a "menu" of at least six schools for every parent to choose from, at least four of which must be of high or medium quality, where quality is determined by test scores. On March 13, 2013 the advisory committee voted to instate Shi's proposal in Boston's Public Schools.

So, with Shi's proposal the final vestiges of busing in Boston disappeared, and the deferred acceptance algorithm was replaced by a restricted choice menu. How does the critical discussion presented in this paper help to understand these developments? BPS was pushed into adopting a different mechanism that performs a different normative theory that *still* fails to offer a democratic forum through which stakeholders might assess whether the algorithm meets the collective demands of the school system.[37] One could run through an inference to best explanation, as presented in subsection 3.2., to locate an ideal theory of justice that corresponds to Shi's algorithm. It would yield a different result—an ideal theory more focused on the tricky problem of walk-zones than the tricky problem of strategizing—but still an elaborate account of what constitutes distributive justice.[38] It is difficult for a stakeholder unhappy with the new system to criticize it on normative grounds, because the ideal theory is enacted. The performative framing of the allocation procedure blocks off normative criticism from participants in the

[36]For a description of the policy conversations between 2005 and 2009, see the *Boston Globe* article Vaznis (2009).

[37]In the final hearing before the committee deliberations, many parents—mostly black and Hispanic parents—expressed concern about Shi's proposal, which was accruing consensus (Seelye 2013a). One father testified "No way we can stand around the playground and say, 'Yeah, we're all getting a fair shake'" (Seelye 2013a).

[38]It might be more along the lines of equal opportunity for good outcomes, where the quality of outcomes is exogenously determined.

optimization problem. The normative gap persists.

Consider the following quotation from a *New York Times* article about Shi's proposal:

> That it took a dispassionate outsider with coding skills but no political agenda to formulate the model is a measure of the complexities facing urban school districts today. Many such districts, like Boston's, are plagued by inequities... Overcoming that legacy here has been so emotionally charged that previous attempts to redraw the zones have failed, though in 2005 the district did change the algorithm it uses to assign students.
>
> (Seelye 2013b)

It is remarkable indeed that the busing system, criticized as "a federal judge's experiment in social engineering," was ultimately axed by a solution drawn up by an MIT graduate student "with coding skills but no political agenda." But I do not take this fact to offer a "measure of the complexities facing urban school districts." Rather, given the critical perspectives I have presented in this section, we can see that it is instead a measure of the gap between an ideal theory enacted through mechanism design and the actual demands of justice. Precisely what Shi and his models did in 2012, and what Abdulkadiroğlu et al. did in 2003, is provide a seemingly objective procedure to "overcome" Boston's legacy of injustice by offering a technology of depoliticization. In this case, to overcome the legacy has been to ignore it by enacting a theory *ab initio* that solves whichever problem happened to have initiated the current policy discussions but not the problems that came before. Millsian criticisms thus arise: How can the mechanism genuinely mete out justice if it fails to explicitly consider certain kinds of actual, present and past injustices in the school system? The normative gap is hard to close through democratic deliberation since each new mechanically objective ideal theory, through the mechanism, is performed. Participants, policymakers and economists alike are trapped in the realm of ideal theorizing.

I do not aim to place blame or responsibility on the economists who have consulted on school choice redesigns. This paper is not directly concerned with the responsibilities of experts in public policy. But this paper does suggest that economists' self-conception as "engineers" is misleading at best and dangerous at worst. Economists do not see themselves as responsible for worrying about "justice" in any deeper way than presenting a mechanism. Their mechanisms have different normative properties that need to be weighed by the policymakers. In some cases the economists clearly express the normative limitations of their designs: "Under the assumptions of the model...the 'leveling the playing field' idea only indicates that sophisticated students lose their strategic rents under the new mechanism" (Pathak 2011: 531). But the economists' technical language on its own is normative in elaborate ways—its machinery operates in the realm of ideal theory. Economists' technology can thus depoliticize sites of democratic contestation, thereby distracting attention from the most important questions at hand.

In this section I have drawn on the concept of performativity to demonstrate what is at stake when ideal theory is enacted through mechanism design. Sen and Mills' criticisms from non-ideal theory demonstrate philosophically that there exists a gap between the ideal theory enacted through the mechanism and the collective demands of the school system. The veneer of objectivity offered by the language of mechanism design helps to explain how the normative gap appeared in the first place. It appeared because the trust of experts had been undermined, and rules and procedures seem to offer an impartial alternative mode of decision-making. BPS took the economists' advice to implement a strategy-proof algorithm not because it addressed the actual issues at hand but because it was a procedure that could be justified to all in a highly abstract realm. The enactment of ideal theory is problematic for the very same reasons that ideal theorizing is problematic to Sen and even Mills. The enactment of ideal theory succeeds

because it offers a performative frame, thereby obstructing stakeholders' avenues for normative criticism of the mechanism.

Conclusion

In the decades that followed *Brown vs. Board of Education,* Boston's school system was subjected to national scrutiny and controversy. Sparking high profile incidents of racially-motivated violence and city-wide protests against court-ordered busing, the allocation of students to Boston's public schools became a lightning rod for profound questions about the nation's approach to racial, social and distributive justice. Three decades after court-ordered busing began in Boston, the mechanism design approach to school choice introduced a novel technical language that would come to mediate Boston's (and other cities') policy discussions in 2005 and beyond. To what extent did this novel technical framing enable policymakers to address inequities in Boston's school system? What is the role of mechanism design theory when used to intervene in policy arenas of enduring social complexity?

In this essay, I offered a critical discussion of the normative forces peculiar to design economics through a close study of several economists' involvement in BPS's policy discussions. I began in section 2 with the social and political background of school choice in Boston. I explained the canonical school choice model from mechanism design (Abdulkadiroğlu and Sönmez 2003), and how BPS came to adopt, in 2005, a version of the famous *deferred acceptance algorithm.* In the following section, I went through an inference to best explanation, suggesting that the adoption of a strategy-proof mechanism instantiated an elaborate normative theory that resembles an *equal opportunity for welfare* principle of justice. In section 4, I argued that the mechanism design thus enacts an *ideal theory.* By considering criticisms of ideal theory (Sen 2011; Mills 2003, 2005), I argued that this enactment of ideal theory inserts a *normative gap* between the abstract ideals of the institution and the collective demands of stakeholders. I drew on the performativity thesis to demonstrate how mechanism design thus serves as a technology of depoliticization.

What does this argument suggest about the future of design economics in policy arenas where issues of social and distributive justice coincide with technological complexities to be overcome? It suggests that the mechanism design approach to school choice will always be flawed and incomplete when instantiated in a particular setting. It reduces families and individuals to calculating individual actors, and flattens complex collective commitments to equality into unarticulated ideal theories of justice. This process closes off democratic debate, as stakeholders are left without the proper technical knowledge and language to voice their opinions about how well the school system is meeting collective demands. Thus, in order for mechanism design to overcome the challenges highlighted in this paper, it must serve as a social technology that invites democratic participation.[39]

Precisely how mechanism design might serve as an inclusive social technology is a pressing topic for future work. Finding a path forward is important for designing school choice procedures in Boston and elsewhere,[40] but also for the design of other justice-oriented institutions. Mechanism design is ripe for application in constructing institutions in wide-reaching realms—from healthcare and town-planning to international trading schemes for carbon and water. If

[39]This imperative relates to views recently advanced to address issues about the use of technology and algorithms in society more broadly—see Crooks (2019) and Hoffmann (2019) for two particularly recent and relevant contributions in this vein.

[40]For an indication of the reach of the theoretical concepts from school choice presented in this paper, consider the reforms of school assignment mechanisms in the U.S. and abroad that have already taken place in the years since 2005. Economic theorists have been involved in designing assignment mechanisms in New York City (Abdulkadiroğlu et al. 2005a, 2009), Chicago (Pathak and Sönmez 2013), Washington DC, Denver and New Orleans (Abdulkadiroğlu et al. 2016; Abdulkadiroglu et al. 2017), India (Bagde et al. 2016) and Brazil (Aygün and Bo 2013). Perhaps most strikingly, a version of the deferred acceptance algorithm has been recently been adopted by all (more than 150) local authorities in England (Pathak and Sönmez 2013).

mechanism design can serve as an inclusive, participative and democratic social technology, it holds the promise of offering tractable solutions to local, national and global problems in which issues of great theoretical complexity intersect explicitly with the real demands of justice, as well as ideals. Economists participating in the market design literature are poised to solve extraordinarily complex problems with many constraints, and a browse through the leading economics journals reveals that market designers are eager to find ways to get their theory to line up with a variety of social goals.[41] But no matter how intricate the mechanism design solutions become, they will necessarily be incomplete as technical solutions to sociotechnical problems. Mechanism designers must strive to develop social technologies that conscientiously enable—rather than accidentally preclude—democratic participation.

Acknowledgements

I am grateful to my supervisor Stephen John for his guidance. I thank Jerry Green, Os Keyes, Scott Kominers, Bobby Pakzad-Hurson, Al Roth, Amartya Sen, Glen Weyl and three anonymous referees for comments and conversations that improved the paper substantially. David Adler, Julian Gewirtz and Indiana Seresin offered valuable responses to earlier drafts. The Frank Knox Memorial Fellowship from Harvard generously supported the writing of this work.

References

Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., and Pathak, P. A. 2016. Charters without lotteries: Testing takeovers in New Orleans and Boston. *American Economic Review*, 106(7):1878–1920.

Abdulkadiroglu, A., Che, Y.-K., Pathak, P. A., Roth, A. E., and Tercieux, O. 2017. Minimizing justified envy in school choice: The design of New Orleans' OneApp. *National Bureau of Economic Research*. Working Paper No. 23265.

Abdulkadiroğlu, A., Pathak, P., Roth, A. E., and Sonmez, T. 2006. Changing the Boston school choice mechanism: Strategy-proofness as equal access. *National Bureau of Economic Research*. Working Paper No. 11965.

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. 2005a. The New York City high school match. *American Economic Review*, 95(2):364–367.

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. 2009. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match. *American Economic Review*, 99(5):1954–1978.

Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. 2005b. The Boston public school match. *American Economic Review*, 95(2):368–371.

Abdulkadiroğlu, A. and Sönmez, T. 2003. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747.

Akbarpour, M. and van Dijk, W. 2018. School choice with unequal outside options. http://web.stanford.edu/ mohamwad/OutsideOptions.pdf.

Alexandrova, A. 2008. Making models count. *Philosophy of Science*, 75(3):383–404.

[41] See for example a recent flourishing of papers around designing for diversity in school choice: Akbarpour and van Dijk (2018); Dur et al. (2018); Erdil and Kumano (2012); Kojima (2012); Budish et al. (2013); Hafalir et al. (2013); Kominers and Sönmez (2016); Echenique and Yenmez (2015).

Alves, M. J. and Willie, C. V. 1987. Controlled choice assignments: A new and more effective approach to school desegregation. *Urban Review*, 19(2):67–88.

Anderson, E. 2010. *The Imperative of Integration*. Princeton: Princeton University Press.

Arneson, R. J. 1989. Equality and equal opportunity for welfare. *Philosophical Studies*, 56(1):77–93.

Arneson, R. J. 1991. A defense of equal opportunity for welfare. *Philosophical Studies*, 62(2):187–195.

Arneson, R. J. 1999. Equality of opportunity for welfare defended and recanted. *Journal of Political Philosophy*, 7(4):488–497.

Austin, J. L. 1955-1962. *How To Do Things with Words*. Oxford: Oxford University Press.

Aygün, O. and Bo, I. 2013. College admission with multidimensional reserves: The Brazilian affirmative action case.

Bagde, S., Epple, D., and Taylor, L. 2016. Does affirmative action work? Caste, gender, college quality, and academic success in India. *American Economic Review*, 106(6):1495–1521.

Barry, B. 1989. *Theories of Justice*, volume 1. Oakland: University of California Press.

Binmore, K. 1994. *Playing Fair: Game Theory and the Social Contract, Volume 1*. Cambridge: MIT Press.

Binmore, K. 1998. *Just Playing: Game Theory and the Social Contract, Volume 2*. Cambridge: MIT Press.

Boldyrev, I. and Svetlova, E., editors 2016. *Enacting Dismal Science: New Perspectives on the Performativity of Economics*. London: Palgrave MacMillan.

Boston Public Schools 2005. Recommendation to implement a new BPS algorithm. (May 11). Available online at http://boston.k12.ma.us/assignment/.

Boston School Committee, Office of the Secretary 1984. School committee secretary desegregation files (1963-1984). City of Boston, Office of the City Clerk, Archives and Records.

Braithwaite, R. 1955. *Theory of Games as a Tool for the Moral Philosopher*. Cambridge: Cambridge University Press.

Breslau, D. 2012. What do market designers do when they design markets? In Camic, C., Gross, N., and Lamont, M., editors, *Social Knowledge in the Making*, chapter 12. Chicago: University of Chicago Press.

Breslau, D. 2013. Designing a market-like entity: Economics in the politics of market formation. *Social Studies of Science*, 43(6):829–851.

Brown v. Board of Education of Topeka 1954. 347 F. Supp. 483 (United States Supreme Court, May 17).

Budish, E., Che, Y.-K., Kojima, F., and Milgrom, P. 2013. Designing random allocation mechanisms: Theory and applications. *The American Economic Review*, 103(2):585–623.

Callon, M. 1998a. The embeddedness of economic markets in economics. *The Sociological Review*, 46(S1):1–57.

Callon, M., editor 1998b. *The Laws of the Markets*. Oxford: Blackwell.

Cook, G. 2003. School assignment flaws detailed: Two economists study problem, offer relief. *Boston Globe*. September 12.

Crooks, R. N. 2019. Times thirty: Access, maintenance, and justice. *Science, Technology, & Human Values*, 44(1):118–142.

Delmont, M. F. 2016. *Why Busing Failed: Race, Media, and the National Resistance to School Desegregation*. Oakland: University of California Press.

Dubins, L. E. and Freedman, D. A. 1981. Machiavelli and the Gale-Shapley algorithm. *The American Mathematical Monthly*, 88(7):485–494.

Dur, U., Kominers, S. D., Pathak, P. A., and Sönmez, T. 2018. Reserve design: Unintended consequences and the demise of boston's walk zones. *Journal of Political Economy*, 126(6):2457–2479.

Dworkin, R. 1981a. What is equality? Part 1: Equality of welfare. *Philosophy & Public Affairs*, 10(3):185–246.

Dworkin, R. 1981b. What is equality? Part 2: Equality of resources. *Philosophy & Public Affairs*, 10(4):283–345.

Echenique, F. and Yenmez, M. B. 2015. How to control controlled school choice. *American Economic Review*, 105(8):2679–2694.

Erdil, A. and Kumano, T. 2012. Prioritizing diversity in school choice. *Unpublished working paper, Washington University*, 1.

Formisano, R. P. 2004. *Boston against Busing: Race, Class, and Ethnicity in the 1960s and 1970s*. Chapel Hill: University of North Carolina Press.

Gale, D. and Shapley, L. S. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

Garfinkel, H. 1956. Conditions of successful degradation ceremonies. *American Journal of Sociology*, 61(5):420–424.

Gauthier, D. 1986. *Morals by Agreement*. Oxford: Oxford University Press.

Guala, F. 2001. Building economic machines: The FCC auctions. *Studies in History and Philosophy of Science Part A*, 32(3):453–477.

Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. 2013. Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363.

Harsanyi, J. 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4):309–321.

Hausman, D. M. 2011. *Preference, value, choice, and welfare*. Cmabridge: Cambridge University Press.

Hausman, D. M. and McPherson, M. S. 2006. *Economic Analysis, Moral Philosophy and Public Policy*. Cambridge University Press.

Hitzig, Z., Hu, L., and Viljoen, S. 2019. The technological politics of mechanism design. *University of Chicago Law Review*, 87(1). Symposium on Radical Markets.

Hoffmann, A. L. 2019. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7):900–915.

Kalai, E. 1977. Proportional solutions to bargaining situations: Intertemporal utility comparisons. *Econometrica*, 45(5):1623–1630.

Kalai, E. and Smorodinsky, M. 1975. Other solutions to Nash's bargaining problem. *Econometrica*, 43(3):513–518.

Kaplow, L. and Shavell, S. 2009. *Fairness versus Welfare*. Cambridge: Harvard University Press.

Kesten, O. 2010. School choice with consent. *Quarterly Journal of Economics*, 125(3):1297–1348.

Kojima, F. 2012. School choice: Impossibilities for affirmative action. *Games and Economic Behavior*, 75(2):685–693.

Kominers, S. D. and Sönmez, T. 2016. Matching with slot-specific priorities: Theory. *Theoretical Economics*, 11(2):683–710.

Landsmark, T. and Dajer, H. 2004. Student assignment task force. *Report and Recommendations of the Boston Public Schools*.

Latour, B. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Li, S. 2017. Ethics and market design. *Oxford Review of Economic Policy*, 33(4):705–720.

MacKenzie, D. A., Muniesa, F., and Siu, L., editors 2007. *Do Economists Make Markets?: On the Performativity of Economics*. Princeton: Princeton University Press.

Mills, C. W. 1997. *The Racial Contract*. Ithaca: Cornell University Press.

Mills, C. W. 2003. *From Class to Race: Essays in White Marxism and Black Radicalism*. Lanham: Rowman & Littlefield Publishers.

Mills, C. W. 2005. 'Ideal theory' as ideology. *Hypatia*, 20(3):165–183.

Mirowski, P., Nik-Khah, E., and MacKenzie, D. 2007. Markets made flesh: Callon, performativity and the FCC spectrum auctions. In *Performativity in the Economic Sciences*. Princeton University Press Princeton, NJ.

Morgan v. Hennigan 1974. 379 F. Supp. 410 (District Court of Massachusetts, June 21).

Naimark, S. 2012. *The Education of a White Parent: Wrestling with Race and Opportunity in the Boston Public Schools*. Amherst: Levellers Press.

Nash, J. 1950. The bargaining problem. *Econometrica*, 18(2):155–162.

Pateman, C. and Mills, C. W. 2007. *Contract and Domination*. Cambridge: Cambridge University Press.

Pathak, P. A. 2011. The mechanism design approach to student assignment. *Annual Review of Economics*, 3(1):513–536.

Pathak, P. A. and Sönmez, T. 2008. Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *American Economic Review*, 98(4):1636–1652.

Pathak, P. A. and Sönmez, T. 2013. School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation. *The American Economic Review*, 103(1):80–106.

Pinch, T. and Swedberg, R., editors 2008. *Living in a Material World: Economic Sociology Meets Science and Technology Studies*. Cambridge: MIT Press.

Porter, T. 1995. *Trust in Numbers*. Princeton: Princeton University Press.

Rawls, J. 1958. Justice as fairness. *The Philosophical Review*, 67(2):164–194.

Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.

Rawls, J. 2001. *Justice as Fairness: A Restatement*. Cambridge: Harvard University Press.

Reid, W. M. 1974. *The Racist Offensive Against Busing: The Lessons of Boston, How to Fight Back*. College Park: Pathfinder Press.

Reiss, J. 2013. *Philosophy of Economics: A Contemporary Introduction*. New York: Routledge.

Richer, M. 1998. Busing's Boston massacre. *Hoover Institute: Policy Review*, (92):42.

Roth, A. E. 1982a. The economics of matching: Stability and incentives. *Mathematics of Operations Research*, 7(4):617–628.

Roth, A. E. 1982b. Incentive compatibility in a market with indivisible goods. *Economics Letters*, 9(2):127–132.

Roth, A. E. 1991. A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom. *American Economic Review*, 81(3):415–440.

Roth, A. E. 2002. The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica*, 70(4):1341–1378.

Roth, A. E. 2007. Repugnance as a constraint on markets. *Journal of Economic Perspectives*, 21(3):37–58.

Schon, D. 1983. *The Reflective Practitioner*. New York: Basic Books.

Seelye, K. 2013a. Boston schools drop last remnant of forced busing. *New York Times*. March 14.

Seelye, K. 2013b. No division required in this school problem. *New York Times*. March 12.

Sen, A. 2006. What do we want from a theory of justice? *The Journal of Philosophy*, 103(5):215–238.

Sen, A. 2011. *The Idea of Justice*. Cambridge: Harvard University Press.

Shapley, L. and Scarf, H. 1974. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37.

Sheehan, J. B. 1984. *The Boston School Integration Dispute: Social Change and Legal Maneuvers*. New York: Columbia University Press.

Shelby, T. 2003. Race and social justice: Rawlsian considerations. *Fordham Law Review*, 72:1697.

Shelby, T. 2013. Racial realities and corrective justice: A reply to Charles Mills. *Critical Philosophy of Race*, 1(2):145–162.

Shi, P. 2015. Guiding school-choice reform through novel applications of operations research. *Interfaces*, 45(2):117–132.

Simmons, A. J. 2010. Ideal and nonideal theory. *Philosophy & Public Affairs*, 38(1):5–36.

Tager, J. 2001. *Boston Riots: Three Centuries of Social Violence*. Boston: Northeastern University Press.

Varian, H. R. 1975. Distributive justice, welfare economics, and the theory of fairness. *Philosophy & Public Affairs*, 4(3):223–247.

Vaznis, J. 2009. Boston shelves 5-zone proposal. *Boston Globe*. August 27.

Biographical Information

Zoë Hitzig is a PhD candidate in economics at Harvard and a Graduate Fellow at the Edmond J. Safra Center for Ethics.