Targeting In-Kind Transfers Through Market Design: A Revealed Preference Analysis of Public Housing Allocation*

Daniel Waldinger†

April 3, 2019

Abstract

Public housing benefits are rationed through waitlists. This paper argues that the range of allocation policies used across U.S. cities involves a trade-off between two policy objectives: maximizing welfare gains for tenants, and targeting the most economically disadvantaged applicants. Using waitlist data from Cambridge, MA, I develop and estimate a model of public housing preferences in a setting where heterogeneous apartments are rationed through waiting time. Counterfactual simulations show that the preferred mechanism depends on social preferences for redistribution. However, many cities use systems that would be suboptimal in Cambridge for any value of redistribution.

JEL Codes: D42, D47, H40
Keywords: Dynamic Matching, Preference Estimation, Discrete Choice Models, In-Kind Transfers

*I am very grateful to Nikhil Agarwal, Parag Pathak, Michael Whinston, and Amy Finkelstein for invaluable guidance and support. I also thank Nick Arnosti, Glenn Ellison, Ingrid Gould-Ellen, John Lazarev, Jacob Leshno, Peng Shi, Neil Thakral, and Heidi Williams; seminar participants at the MIT Industrial Organization and Public Finance workshops; my colleagues at NYU; and participants in many other seminars and conferences. The Cambridge Housing Authority generously provided the applicant and tenant data used in this paper, with special thanks to Tara Aubuchon, Tito Evora, Michael Johnston, Jay Leslie, Hannah Lodi, and John Ziniewicz. All analysis and views expressed in this paper are my own and do not represent the views of the Cambridge Housing Authority. I acknowledge support from a National Science Foundation Graduate Research Fellowship. All mistakes are my own.

†NYU Furman Center. Email: danielwaldinger@nyu.edu.

1 Introduction

In the United States, more than one million low-income households live in public housing. Beneficiaries receive a permanent, place-based entitlement to a rent subsidy that often exceeds \$10,000 per year. However, due to limited funding, assistance is rationed – in 2012, there were at least 1.6 million additional households on public housing waitlists nationwide (Collinson et al., 2015). Given this scarcity, public housing authorities (PHAs) in each city use a wide variety of rules to allocate available apartments. These systems differ dramatically in the degree of choice applicants have over where they can live, and also which applicants are prioritized to receive apartments first.

This paper empirically studies the consequences of these rationing mechanisms. I focus on two policy objectives: maximizing tenants' values of their assignments, and targeting assistance to the most economically disadvantaged applicants. Theoretical work has shown that there can be a trade-off between these two goals, but its existence and economic importance are not well understood empirically (Arnosti and Shi, 2018). Since public housing is a large anti-poverty program, distributional concerns are especially relevant. Many households without housing assistance are extremely low-income, facing high rent burdens and eviction rates on the private rental market (Collinson and Reed, 2018; Desmond, 2012; Humphries et al., 2018). It is therefore crucial to understand how public housing rationing mechanisms affect *who* receives assistance, in addition to how much beneficiaries value it.

To answer this question, this paper develops an empirical strategy to evaluate public housing allocation policies and applies it to waitlist data from the Cambridge Housing Authority (henceforth, CHA), which administers public housing in Cambridge, MA. The first part of the analysis recovers the distribution of preferences for public housing developments based on applicants' decisions while on the waitlist. A central idea is that waiting time acts as a price – some applicants face a trade-off between waiting for less time and being assigned to their preferred housing development. The second part of the analysis uses the estimated model to predict equilibrium allocations if the CHA adopted other cities' allocation mechanisms. This step requires solving for new equilibrium waiting time distributions that are consistent with applicants' optimal decisions.

I find that existing policies involve a dramatic trade-off between efficiency – the welfare gains generated by assignments – and redistribution – the fraction of extremely high-need tenants. This trade-off is driven by substantial heterogeneity in both applicants' preferred housing developments and their overall values of obtaining assistance. However, despite the importance of this trade-off, there are combinations of choice and priority which are dominated in a broad class of social welfare functions. These mechanisms are not optimal in Cambridge for any social preference for redistribution.

To gain intuition for how waitlist design can influence which applicants receive assistance, imagine applying for public housing in New York City and Miami. In New York City, you are asked to choose your preferred housing development and wait until an appropriate apartment becomes available. In

Miami, you are offered the first available apartment from any development when you reach the top of the waitlist, and removed from the list if you decline. These development choice rules give you very different incentives. New York City allows you to wait for your preferred development, while Miami only allows you to express your preferences by declining assistance entirely. Theoretical work has argued that not allowing choice, as in Miami, will often lead to lower *match quality* for tenants because applicants must accept mismatched offers in order to be housed (Bloch and Cantala, 2017; Leshno, 2017). However, not allowing choice may improve *targeting* if applicants with better outside options are more likely to reject, self-selecting out of the application process and allowing more economically disadvantaged applicants to be housed. The trade-off depends crucially on applicant preferences – where applicants want to live, how much they value assistance compared to their outside options, and sensitivity to waiting time.

A key contribution of the first part of this paper is to develop a method to estimate applicant preferences using waitlist data. To my knowledge, this type of data has not yet been used to estimate demand for public housing. Here, waitlist data address an important empirical challenge: because public housing is rationed through a waitlist, one cannot infer that eligible households not living in public housing do not desire to, or that public housing tenants live in their preferred developments. The CHA data record which households applied for CHA housing over a 5-year period and contain rich development choice information, providing a direct measure of which households expressed demand for assistance as well as willingness to wait for specific developments.

To recover the distribution of preferences from applicant decisions, I build a structural model of application and development choice based on the structure of the CHA's allocation mechanism, henceforth the Cambridge Mechanism. In the Cambridge Mechanism, applicants choose their preferred housing development in two stages, and receive new information about waiting time in the second stage. Given this structure, an applicant's development choice problem is a generalized version of the portfolio choice problem considered in theoretical work by Chade and Smith (2006). I provide descriptive evidence that second-stage choices are responsive to new information, suggesting that responsiveness to waiting time is an important consideration when evaluating alternative allocation mechanisms. I therefore propose a development choice model in which applicants know their preferences over housing developments but face waiting time uncertainty. They choose their preferred *distributions* of assignments and waiting times at each stage of the application process, given their current information.

The estimation procedure involves three steps which build on techniques from the literature on demand estimation and empirical market design. First, to account for the fact that applicants may be a selected sample of eligible households, I estimate the distribution of households eligible for CHA public housing during the period of study using American Community Survey data. Second, I construct applicants' beliefs about the distributions of assignments and waiting times they face. I assume that beliefs match a long-run steady state of the Cambridge Mechanism generated by the observed distribution of applicants and their decisions. This assumption allows me to construct beliefs from the output of a detailed simulation of the Cambridge Mechanism, using the data to estimate a lower-dimensional set of inputs. Finally, I estimate preference parameters by matching predicted choice patterns to those in the data using the method of simulated moments (McFadden, 1989; Pakes and Pollard, 1989).

I parameterize applicant preferences in a way that highlights the potential trade-off between efficiency and redistribution. Applicants have a common discount factor, but heterogeneous preferences both for specific developments (*match values*) and for the aspects of public housing that are common across developments (*values of assistance*). The value of assistance is interpreted relative to an applicant's outside option. To motivate this decomposition and facilitate welfare analysis, I assume a utility model in which applicants have heterogeneous tastes for specific public housing developments and different levels of income outside of public housing. Importantly, part of the value of the household's outside option is not observed but affects their marginal utility of income. With restrictions on the functional form of utility, indirect utilities can be converted to equivalent variation.

The application data and structure of the Cambridge Mechanism provide crucial information about both dimensions of preference heterogeneity. Application rates by income and demographic groups are informative about differences in values of assistance. For example, lower-income and non-white households are more likely to apply than other eligible households, suggesting they have higher values of assistance on average. Initial development choices are informative about heterogeneity in match values and values of assistance. Since CHA applicants choose up to three lists in the first stage, these initial choices reveal not only which developments are more likely to be chosen overall, but also which developments tend to be chosen by the same applicant. These patterns reveal heterogeneity in tastes for specific development characteristics, including waiting time, as well as idiosyncratic tastes. Sensitivity of second-stage development choices to waiting time information allows me to estimate the discount factor in addition to the parameters governing flow payoffs. The moments used in estimation capture these features of application decisions and development choices.

Estimates show that applicants are patient and exhibit substantial heterogeneity in values of assistance and match values. While observed characteristics strongly predict the value of assistance, particularly income and race, the standard deviation of the unobservable among applicants is equivalent to more than $13,000 of annual income. Applicants also have strong preferences for specific developments. For the median applicant, the difference in value between living in their first and second choice developments is equivalent to a cash transfer of $563 per year. Due to strong development preferences, some applicants are quite selective. 15 percent of applicants would only be willing to live in three or fewer developments, while half of applicants prefer any development to their outside option. The latter group have much lower observed incomes and unobservably worse outside options. Therefore, a mechanism designer may be able to increase redistribution by inducing rejections.

In the second part of the paper, I consider how the development choice and priority systems used by other cities would affect the equilibrium allocation of public housing in Cambridge. These counterfactuals only change allocation policy, holding fixed the Cambridge public housing stock and other market primitives. For each policy, I solve for a fixed point between applicants' optimal decisions and their endogenously generated waiting time distributions. I focus on existing policies partly due to the absence of applicable theoretical results on optimal dynamic mechanisms, but also because such diverse mechanisms are used in practice. In addition to giving applicants different amounts of choice, cities differ in whether higher- or lower-income applicants receive priority. Thus, I can explore a wide range of outcomes, in terms of efficiency and redistribution, under mechanisms cities already use.

Under the current CHA priority system, a move from allowing choice to no choice redistributes housing towards applicants with lower incomes at the cost of worse matches. Mean tenant incomes fall from \$18,252 to \$16,903; meanwhile, the fraction of tenants living in their first choice development falls from 52 to 8 percent, and equivalent variation per assigned unit falls by more than 30 percent. In contrast, the effects of prioritizing higher- or lower-income applicants are mainly distributional: holding the choice system fixed, cost-adjusted welfare gains and match quality are similar across priority systems. These results suggest that PHAs can increase redistribution through the priority system without sacrificing efficiency.[1]

To conclude the analysis, I ask how distributional preferences determine which mechanism the CHA should use. I consider a social planner who prefers making transfers to more economically disadvantaged applicants, using a "constant relative inequality aversion" class of social welfare functions (Atkinson, 1970) to compare allocation mechanisms across a range of values of redistribution. To assign social welfare weights to applicants, I assume that conditional on observed characteristics, applicants with higher values of assistance have higher marginal utilities of income.

This exercise shows that removing choice should be a policy of last resort to improve targeting once observed characteristics have been exploited through the priority system. Intuitively, prioritizing lower-income applicants and removing choice both improve targeting, but removing choice has a greater efficiency cost due to lower match quality. When inequality aversion is moderate, it is best to prioritize lower-income applicants while allowing choice; only when inequality aversion is very high is it optimal to remove choice in addition to prioritizing lower-income applicants, maximizing redistribution. A corollary of this insight is that certain combinations of choice and priority are dominated for any degree of inequality aversion. In particular, it is never optimal to prioritize higher-income applicants while not allowing choice. This mechanism is used in Los Angeles, but it would not perform well in Cambridge – there is a better policy whether society has a high or a low value of redistribution.

The paper proceeds as follows. Section 1.1 discusses this paper's contribution to the literature.

[1] An important caveat to these results is that the counterfactual simulations assume that households will not manipulate their characteristics in response to income-based priorities. To my knowledge, there is no evidence on endogenous responses to changes in a priority system in a dynamic matching mechanism, and the data and setting here were not suitable to provide evidence. It is worth noting that households would have to manipulate their incomes over a long period of time.

Section 2 provides institutional background on the U.S. public housing program, discusses trade-offs involved in allocation policy, and describes the CHA dataset. Section 3 presents descriptive facts about Cambridge public housing developments, applicants, and their decisions while on the waitlist. Motivated by these patterns, Section 4 proposes a model of development choice as well as a micro-foundation of preferences used for welfare analysis. Section 5 describes the parametric model and the estimation procedure used to recover the distribution of preferences for public housing developments. Section 6 presents the estimation results, and Section 7 analyzes the effects of alternative choice and priority systems. Section 8 concludes.

1.1 Related Literature

This paper contributes to several literatures on centralized matching markets, means-tested housing assistance, and in-kind transfers.

This paper joins a growing literature on revealed preference analysis in centralized matching markets (Abdulkadirolu et al., 2017; Agarwal, 2015; Agarwal and Somaini, 2018; Fack et al., 2019; Hastings et al., 2009; He, 2017; Narita, 2018). Much of this work has focused on static problems. Along with two recent papers, this paper is among the first to conduct revealed preference analysis using the decisions of agents in a dynamic assignment mechanism. Agarwal et al. (2019) study the allocation of deceased donor kidneys, where patients face an optimal stopping problem in deciding which organs to accept. Reeling and Verdier (2018) study the allocation of hunting licenses in Michigan, finding that allowing choice improves both match quality and targeting because hunting sites are highly vertically differentiated. Both of these papers employ techniques from dynamic discrete choice estimation, whereas the present work analyzes a portfolio choice problem drawing on techniques from static discrete choice. This paper is also unique in considering income redistribution as an explicit policy objective in the design of a centralized matching market. Dworczak et al. (2018) theoretically study how the goal of income redistribution influences the optimal design of a decentralized market when, as in my application, the tax schedule and policies in other markets cannot be adjusted.

The empirical papers most closely related to this work estimate demand for public housing using data on assignments (Geyer and Sieg, 2013; Sieg and Yoon, 2016; Van Ommeren and Van der Vlist, 2016). To my knowledge, this paper is the first to use individual-level waitlist data to estimate demand for public housing. Other empirical work has argued that there is substantial misallocation in the public and rent-controlled housing sectors (Glaeser and Luttmer, 2003; Thakral, 2016). Consistent with this work, I find that public housing allocation policy can have a large impact on match quality for tenants. A complementary literature evaluates the causal effects of housing assistance, and has found that receiving housing assistance and living in higher socioeconomic status neighborhoods can lead to improved economic and health outcomes (Andersson et al., 2016; Chetty et al., 2015; Kling et al., 2007; Ludwig et al., 2013; Van Dijk, 2018). The subjective values for public housing estimated in this paper may include households' beliefs about the program's long-term benefits in addition to immediate changes in disposable income and housing and neighborhood quality.

The market design trade-off between match quality and targeting has been studied in the theoretical literature on one-sided dynamic assignment (Arnosti and Shi, 2018; Bloch and Cantala, 2017; Leshno, 2017; Thakral, 2016). Arnosti and Shi (2018) show that the relationship between match quality and total welfare is theoretically ambiguous and depends on the distribution of agent preferences. This paper provides empirical evidence on these primitives and their implications for allocation policy. A related literature on the targeting of public assistance has highlighted the tension between providing valuable assistance to those who receive it ("productive efficiency") and restricting assistance to the households which need it most ("targeting efficiency") (Akerlof, 1978; Nichols and Zeckhauser, 1982). Several recent papers have studied this idea empirically in the context of means-tested transfer programs of homogeneous benefits (Alatas et al., 2016; Deshpande and Li, 2019; Finkelstein and Notowidigdo, 2018; Leiber and Lockwood, 2019). This paper explores an analogous trade-off that arises for very different reasons: because public housing is heterogeneous and in limited supply, a market designer can lower expected match quality to screen out low-value applicants.[2] I also analyze how applicant priorities, a version of "tagging" (Akerlof, 1978), interact with alternative development choice systems.

2 Institutional Background and Data

Section 2.1 provides an overview of the U.S. public housing program, surveys allocation policies used in practice, and discusses the design trade-offs these policies entail. Section 2.2 describes the Cambridge Housing Authority and the mechanism it used to allocate public housing during the period of study. Section 2.3 describes the applicant dataset and sample criteria.

2.1 Public Housing in the U.S.

The U.S. public housing program subsidizes the rents of 1.2 million low-income households at an annual cost of $8-10 billion. A Public Housing Authority (PHA) in each city maintains the stock of public housing developments located in its jurisdiction using funds allocated by Congress and distributed by the U.S. Department of Housing and Urban Development (HUD). A public housing tenant pays 30 percent of pre-tax income toward rent, and is permanently entitled to assistance as long as it complies with the terms of its lease and remains in its assigned apartment. Public housing and its private market counterpart, the Housing Choice Voucher program, are unusual in their benefit generosity: in 2013,

[2]The fact that public housing involves an in-kind transfer of housing rather than cash may also sacrifice productive efficiency by distorting the housing consumption of those who receive assistance. Given that only one quarter of eligible households applied for Cambridge public housing during the period of study, the targeting gains from public housing may be large compared to a cash transfer of equal value. Though the data and institutional setting in this paper are not suited to answer the question of in-kind versus cash assistance, it is an important question for future research.

participants received an average annual subsidy of $8,000.[3]

Due to the combination of limited federal funding, generous per-household benefits, and broad eligibility criteria, demand for public housing greatly exceeds supply. Congress does not set funding levels to assist all eligible households, but rather to maintain existing services. New public housing is not being built.[4] The income limit for eligibility is 80 percent of Area Median Income (AMI), a regional income measure adjusted for household size, which includes lower-middle income households as well as the very poorest. As a result, in 2012 there were approximately 1.6 million households on public housing waitlists nationwide, and nearly 3 million applicants on voucher waitlists.[5]

2.1.1 Public Housing Allocation Mechanisms and Design Trade-Offs

The limited supply of public housing creates a dynamic assignment problem for each PHA. When tenants move out, the PHA must assign vacant apartments to applicants on a waitlist. PHAs have substantial autonomy over allocation policy. In particular, they control how applicants are ordered on the waitlist and whether applicants can choose the developments to which they are assigned. These policy levers – the *priority system* and *development choice system* – can affect which types of applicants receive assistance and whether they are matched to their preferred developments. To my knowledge, there is no resource that systematically documents the current waitlist policies of each of the 3,300 U.S. PHAs. To summarize allocation policies used in practice, I examined most recent available administrative plans of 24 PHAs falling into two categories: (1) those with the largest public housing stocks, and (2) those with public housing stocks and city populations similar to Cambridge, MA. The priority and development choice systems used by these PHAs are summarized in Table 1.

The allocation policies of surveyed PHAs share several common features. The lists are first come first served; applicants are ordered by priority and then by date of application. If an applicant rejects an offer without good cause, they are removed from the list, though some PHAs allow one or two rejections prior to removal. PHAs offer apartments to applicants living or working in the jurisdiction before other applicants, and there are also federally mandated priorities for specific groups, such as veterans and natural disaster victims.

Despite these similarities, existing choice and priority systems exhibit important differences. The key difference across priority systems is whether households with higher or lower socioeconomic status receive priority. Some PHAs, including those in New York City and Los Angeles, give priority to

[3] Based on per-household subsidy from tenant-based vouchers reported in HUD Congressional Justification for FY2015, available at `https://www.hud.gov/sites/documents/FY15CJ_PUB_HSNG_CAPTL_FND.PDF`. In 2013, the public housing program served a population with similar incomes.

[4] The majority of new affordable housing is built through the Low-Income Housing Tax Credit (LIHTC), a federal tax expenditure that subsidizes private sector construction of new affordable housing. This program is administratively separate from the public housing and voucher programs and has a different rent payment structure, so that tenants with very low incomes receive a smaller effective rent subsidy than in public housing.

[5] Public and Affordable Housing Research Corporation (PAHRC), 2015. "Value of Home: 2015 PAHRC Report." Based on PAHRC tabulation of the Public Housing Agency Homelessness Preferences Survey, 2012.

households with a working member, that are economically self-sufficient, or that have incomes above 30 percent of AMI. Others do just the opposite – the Seattle Housing Authority prioritizes households below 30 percent of AMI, and several other PHAs prioritize households that are severely rent burdened or at risk of being displaced. To the extent that observed characteristics capture the underlying characteristics of interest – for example, some notion of "need" – priorities can effectively target desired applicants. However, if prioritized characteristics only weakly predict need, priorities can actually make targeting worse by shutting out a large fraction of high-need applicants. It may be preferable to induce applicants to reveal residual private information through their decisions on the waitlist.

The range of development choice systems across PHAs is equally wide. Several PHAs, including those in New York City, Seattle, and New Haven as well as Cambridge, require applicants to choose a limited number of developments (Limited Choice). As noted in the dynamic market design literature, Limited Choice tends to achieve good match quality because applicants with the highest values of over-subscribed developments will be more likely to apply for and occupy them. Other PHAs do not allow applicants any choice over their assignment (No Choice), including large cities such as Miami, Los Angeles, and Minneapolis. Two conditions must be met for No Choice to improve targeting. First, some applicants must be willing to reject undesirable apartment offers; second, the social value of housing these selective applicants must be lower than for applicants who are less likely to reject. PHAs also use intermediate choice systems. Chicago allows applicants to select a neighborhood but not a specific development, while in Boston, applicants may choose any subset of developments, providing the option to hedge against waiting time uncertainty. Under some conditions, these systems will produce intermediate allocations in terms of match quality and targeting (Arnosti and Shi, 2018).

PHAs combine development choice and priority systems in different ways. Los Angeles uses No Choice, but prioritizes applicants that are economically self-sufficient (high-SES). Seattle does the reverse, allowing Limited Choice while prioritizing low-SES applicants. Minneapolis uses both development choice (No Choice) and priorities (low-SES) to maximize targeting, while New Haven prioritizes higher-income applicants and allows choice. I will argue that in Cambridge, only certain combinations of these development choice and priority systems could be justified by preferences for income redistribution.

2.2 The Cambridge Housing Authority

The Cambridge Housing Authority (henceforth, CHA) administers the Public Housing and Housing Choice Voucher programs in Cambridge, MA. During the period of study – January 1st, 2010 to December 31st, 2014 – the CHA public housing stock consisted of about 2,450 apartments, evenly split between the Elderly/Disabled and Family Public Housing programs. Although Cambridge is a low-poverty area compared to a nationally representative sample of public housing sites, Cambridge public housing tenants are comparable to those nationwide in terms of socioeconomic status and demographics.

In 2014, 74 percent of Cambridge public housing tenants earned less than 30 percent AMI and 48 percent were headed by an African American, compared to 72 percent and 48 percent nationwide.

The CHA employed a site-based waitlist system to allocate public housing during the sample period. The waitlist for vouchers was closed from 2008 until 2016, while public housing waitlists were open from 2008 until 2015. For this reason, I study the public housing program in isolation.[6]

2.2.1 The Cambridge Mechanism

In the Cambridge Mechanism, applicants select their preferred development in a two-stage process. Each development is a building, complex, or collection of apartments in a distinct geographic location, and apartments with the same number of bedrooms are mostly homogeneous within a development. All applicants with a household member living or working in Cambridge receive equal priority. The development choice system shares features with those used in New York City, Seattle, and New Haven; the priority system is similar to those used in Chicago, Philadelphia, and Boston.

A key difference between the Cambridge Mechanism and many other development choice systems is that applicants choose their preferred development in two stages.[7] At initial application, a household is assigned a program (Elderly/Disabled or Family) and bedroom size and makes an initial choice of up to three developments from 9 to 13 alternatives. The initial choice forms the applicant's choice set later on, and the applicant is placed on a waitlist for each chosen development. At a later date, the CHA sends the applicant a letter asking them to make a final development choice. The letter informs the applicant of their current position on each list in their choice set, which provides new information about continued waiting time. After making its final choice, the applicant remains on the waitlist for that development until the CHA makes a single, take-it-or-leave-it offer of an apartment. If the applicant rejects, it is removed from the waitlist and cannot reapply for one year. The applicant may also be removed if it fails to attend its screening appointment, produce required documentation, or respond to mail from the CHA. Appendix C.1 provides a formal description of the Cambridge Mechanism, including when the CHA sends final choice letters and how it calculates list position.

2.3 Dataset and Sample Selection

The main dataset used in this paper, provided by the CHA, contains anonymized records of all applicants for Cambridge public housing who were active on a waitlist between October 1st, 2009 and February 26th, 2016. The CHA maintains a database of applicants to manage its waitlists and comply with HUD regulations. For each applicant, the dataset records household characteristics, development choices, and the timing and outcome of all events during the application process.

[6] Every year, each housing authority is required to publish an Admissions and Continued Occupancy Policy (ACOP). The following description is based on the CHA's 2014 ACOP as well as conversations with several CHA employees.

[7] The New York City Housing Authority, which administers 15 percent of the nation's public housing stock, uses a similar two-stage development choice system. Applicants first choose a preferred borough, and later choose their preferred development from a subset of the developments in that borough.

For analysis, I restrict my sample to applicants who had priority for Cambridge public housing; who applied for 2 and 3 bedroom apartments in the Family Public Housing program; and who submitted an application between 2010 and 2014. Non-priority applicants had virtually no chance of being housed, and are therefore excluded. Family Public Housing applicants are a more homogeneous group than Elderly/Disabled applicants. I restrict to 2 and 3 bedroom apartments for sample size; within family public housing, there are few apartments and applicants for other bedroom sizes. Analyzing new applications between 2010 and 2014 avoids selection issues because not all pre-2010 applicants were still on the waitlist in 2010. These restrictions produce a sample of 1,752 applicants. After omitting 26 irregular applications, 1,726 applicants remain.

To estimate the distribution of households which *could* have applied during the sample period, I augment the CHA dataset with a sample of eligible households from the American Community Survey (ACS). I also use data provided by the CHA on Cambridge public housing tenants between 2012 and 2014. Appendix A provides details of the CHA and ACS datasets, and Section 5.1 explains how they are used to estimate the distribution of potential applicants.

3 Descriptive Evidence

This section presents descriptive statistics describing Cambridge public housing applicants and their development choices. These facts illustrate the key economic forces that will be quantified in the structural model. Cambridge public housing developments differ in size, location, and expected waiting time. Applicants' decisions to apply and initial development choices reveal heterogeneity in their overall values of living in public housing, as well as where they want to live. While observed characteristics strongly predict who applies and which developments they prefer, much choice behavior is left unexplained. Final choices reveal that applicants are sensitive to waiting time information, and will choose a less preferred development in exchange for a shorter expected waiting time.

3.1 Cambridge Public Housing Developments

During the period of study, applicants for Family Public Housing in Cambridge chose among thirteen developments located throughout the city. The location of each development is shown in Figure 5. There are three developments in East Cambridge, three in North Cambridge, and seven near Central Square. Table 3 displays characteristics of these developments. The smallest developments contain just a few apartments that blend in with the surrounding housing stock,[8] while the largest developments are complexes of several buildings containing hundreds of apartments. Developments also have different expected waiting times. Average waiting times for housed applicants range from 1.58 to 3.75 years

[8] The "Scattered" waitlist represents three lists: one for scattered sites in Mid-Cambridge (Central), one for East-Cambridge, and one for River Howard Homes (Central). In July 2013, the CHA combined Mid Cambridge, River Howard Homes, and Woodrow Wilson with Putnam Gardens, and also combined East Cambridge with Roosevelt Low-Rise. Only Putnam Gardens and Roosevelt Low-Rise were options thereafter, reflecting units from the combined lists.

across developments, with smaller developments tending to have longer waits. As a result, some applicants faced a trade-off between their preferred assignment and a shorter expected wait. Developments are less heterogeneous in terms the characteristics of their tenants, with similar average incomes and proportions of African American tenants.[9]

3.2 Application Decisions and Initial Development Choices

Application rates by income and demographic groups reveal which types of households value public housing the most relative to their outside options. The first two columns of Table 2 show that only one in four eligible households actually applied for Cambridge public housing during the sample period. Those who did apply had much lower incomes and were more likely to be non-white and to already live in Cambridge. The average income of eligible households is $42,219, while that of applicants is $18,477. This is to be expected; since rent is 30 percent of pre-tax income, a lower-income household sees larger increases in housing quality and disposable income in public housing compared to its outside option. Differences by race are also striking: half of applicant households are headed by an African American, while only one in five eligible households are. Although income and race strongly predict who applies, they are not perfectly predictive. Figure 6 shows that while application rates fall steadily as income rises, some of the lowest-income households did not apply and some high-income households did. Similarly, 25 percent of African American headed households did not apply.

The remaining columns of Table 2 show that most applicant characteristics are stable over time. However, there are a couple of moderate trends. The rate of new applications fell from 415 per year in 2011 to 347 in 2014.[10] Over time, new applicants had higher incomes and were more likely to work in Cambridge and have a white head of household. Applicant income growth is consistent with improving economic conditions in Massachusetts following the Great Recession. Despite the fact that only one in four eligible households applied for public housing during the sample period, there were five applicants for each of the 327 apartment vacancies. Demand greatly exceeded supply in this market.[11]

Initial development choices suggest that applicants have strong tastes for specific developments and that their preferences are correlated with observed characteristics. Table 4 presents statistics from initial development choices for all applicants and broken out by household income and neighborhood of current residence. Applicants that already live in Cambridge are much more likely to select developments in their own neighborhoods. The majority of applicants (84 percent) exhaust their initial choice set and select three housing developments. This rate is lower for applicants with incomes over $32,000:

[9] There are outliers. For example, Roosevelt Mid-Rise has an unusually low average tenant income and a small fraction of African American tenants. This is because it is a mixed development, with some apartments for Elderly and Disabled households. Its tenants are older, and as a result have lower incomes and are more likely to be white.

[10] The CHA closed its Family Public Housing waitlists during the second and third quarters of 2010. As a result, 2010 saw fewer new applications than subsequent years.

[11] The number of vacancies is below the long-run average because the CHA began renovating its public housing stock during the sample period. For a plausible upper bound on the long-run average, an annual turnover rate of 10 percent per unit would raise the expected number of vacancies to 540 over a five year period.

only 78 percent select three lists, compared to 85 percent for lower-income applicants. Higher-income applicants also select developments with slightly longer average waiting times. These patterns are consistent with a model in which applicants with better outside options are more selective in their development choices. However, the fact that these differences are not larger suggests the presence of unobserved heterogeneity in values of assistance.[12] Similarly, specific chosen developments are not fully predicted by observed characteristics. The structural model will quantify heterogeneity in both values of assistance and match values, as a function of both observed and unobserved characteristics.

3.3 Response to Waiting Time Information

This section presents quasi-experimental evidence that applicant choices are sensitive to information about waiting time. Between 2010 and 2014, the CHA sent final choice letters to applicants who were near the top of the list for one of their initial choice developments. The letter informed applicants of their position on each list and asked them to make a final development choice. Because of fluctuations in relative list lengths over time, and also due to the CHA's algorithm for calculating list position and sending final choice letters, applicants who made the same initial development choices but applied on different dates were given different position information when they made their final choices. Final choices are sensitive to this information: when an applicant is told a lower list position for one development relative to the others in their choice set, they are more likely to pick that development.

To test the null hypothesis of no response to waiting time information, I run a conditional logistic regression that predicts an applicant's final choice as a function of list position or expected continued waiting time. The sample is applicants who made a final choice during the period of study, and the outcome is which development they chose. Since each applicant chose their choice set at initial application, I include as controls fixed effects for the interaction between each development and choice set. This isolates the natural experiment in which applicants who made the same initial choices – and whose development preferences are therefore drawn from the same distribution – receive different waiting time information for the same alternatives. In this specification, choice probabilities are

$$P(j \mid C_i) = \frac{\exp\{\beta x_{ij} + \xi_{j,C_i}\}}{\sum_{k \in C_i} \exp\{\beta x_{ik} + \xi_{k,C_i}\}}, \tag{1}$$

where C_i is applicant i's set of initially chosen developments, conditioning on bedroom size; x_{ij} is the position number or expected continued waiting time for development j; and ξ_{j,C_i} is a fixed effect for the interaction between development j and initial choice C_i.

Table 5 displays coefficient estimates and implied marginal effects from specifications with no controls; with development fixed effects; and with the full set of development and choice set interactions

[12] Note that higher-income households who applied for Cambridge public housing are already a selected sample. This should mute any correlation between applicant characteristics and the selectivity of their development choices.

in equation 1. For each set of controls, the specification is run for both list position and expected continued waiting time. Except for column (2), coefficient estimates are precise and show a negative response to list position and continued waiting time. The response grows stronger with additional controls, and the implied elasticities are large: with full controls, the elasticity of final choice is -1.1 with respect to list position and -4.1 with respect to continued waiting time.

For a test of the null hypothesis of no response to be valid, position information must be uncorrelated with development preferences among applicants with the same choice set who made a final choice. Two conditions are sufficient for this assumption to hold. The first is that the development preferences of applicants who applied on different dates but made the same initial choice are drawn from the same distribution. This would not be true if, for example, applicants anticipated fluctuations in waiting times, since this would influence initial choices. However, given that waiting time fluctuations are determined by randomness in when apartments become vacant and the decisions of other applicants, these fluctuations would have been difficult to predict or influence. The second condition is that response to the final choice letter is uncorrelated with the specific information in the letter, conditional on the elapsed time since application. This will be true if applicants become unresponsive for exogenous reasons. While these conditions are difficult to test directly, Appendix B shows that applicants' initial choices are not predicted by list lengths on the specific date they applied, and that their observable characteristics are not predicted by the information they receive in the final choice stage.

These results simply establish the existence of a response. In structural estimation, moments based on responsiveness to waiting time information will separate applicants' rate of time preference from heterogeneity in their values of specific developments.

4 Model of Preferences and Development Choice

To interpret applicants' decisions while on the waitlist and predict equilibrium behavior and welfare under alternative allocation mechanisms, this section develops a model of applicant preferences and development choice. Section 4.1 presents a development choice model which predicts how eligible households behave at each stage of the Cambridge Mechanism given their preferences and information. This model is sufficient for positive analysis of alternative waitlist policies. Section 4.2 provides a micro-foundation of preferences that links development preferences to households' marginal utilities of income. I use this model for normative analysis quantifying welfare and distributional impacts.

4.1 Development Choice Model

The development choice model provides a rational benchmark through which to interpret the application decisions of eligible households and development choices of applicants. In particular, it captures the trade-off between spending less time on the waitlist and being assigned to a preferred housing development. Since alternative allocation mechanisms will lead to different equilibrium waiting time distributions, quantifying how applicants make this dynamic trade-off is crucial for predicting equilibrium allocations.

In the model, an applicant solves a single-agent problem, taking the strategies of other applicants as given. She receives flow indirect utility depending on where she lives – outside of public housing, or in a specific public housing development – and discounts future payoffs. These preferences do not change over time, but the applicant does make forward-looking decisions. The Cambridge Mechanism generates a portfolio choice problem because applicants make development choices in two stages and receive new information in the second stage. The applicant understands the structure of the Cambridge Mechanism and solves the portfolio choice problem correctly. She faces waiting time uncertainty because she does not know exactly who is on the waitlist or how it will evolve in the future. Therefore, she forms beliefs about the joint distribution of assignments and waiting times each decision will induce, and chooses her preferred distribution at each stage of the application process. In estimation, I place additional structure on beliefs and assume that applicants have a common prior when making their initial choice, and update their beliefs based on the information provided at the final choice stage.

The following sections specify the sequence of decisions; payoffs; information and beliefs about how choices affect future states; and the resulting portfolio choice problem.

4.1.1 Sequence and Timing of Decisions

An eligible household, indexed by i, makes decisions in the following sequence:

1. *Application Decision*: i receives the opportunity to apply on a random date.
2. *Initial Choice*: If i applies, she immediately chooses up to three developments, denoted $C \subset \{1, ..., J\}$ with $|C| \leq 3$. These developments form i's choice set in the final choice stage, and i is placed on a waitlist for each development $i \in C$.
3. *Final Choice*: At a later date, i receives a letter containing her position on the waitlist for each development in her choice set. The letter asks i to make a final choice $f \in C$. Let s denote the number of years between initial application and the final choice letter, and let $p \equiv \{p_j\}_{j \in C}$ denote the vector of list positions. If i responds to the letter and chooses development f, she remains on the waitlist until she receives a take-it-or-leave-it apartment offer in f.[13]

Applicant i may become unresponsive at any point during the application process and is removed from the waitlist if this occurs. I assume that attrition is exogenous to the model; that an applicant cannot anticipate the date she will be removed; and that removal occurs at a poisson rate α that is equal across applicants. Applicants may not fully anticipate the possibility of attrition, and have a

[13] In principle, the applicant also decides whether to accept their offer of housing. In the model presented here, an applicant will always accept an offer from the Cambridge Mechanism. This is because they would never choose a development that is worse than their outside option in the final choice stage, and apartments are homogeneous within a development. Rejections occurred infrequently in the CHA data.

subjective attrition probability $\tilde{\alpha} \leq \alpha$. This allows an applicant's effective discount rate to be lower than their rate of attrition.

4.1.2 Preferences over Assignments and Waiting Times

Household i receives a payoff that is realized continuously over time and depends on where she lives. In particular, i's per-period flow indirect utility from living in development j is v_{ij}, and her flow indirect utility from not living in Cambridge public housing is v_{i0}. I will refer to these flow indirect utilities as flow payoffs. Assignments are believed to be permanent, and anticipated flow payoffs are not time-dependent. This rules out learning about characteristics of the developments over time or changing household circumstances. When making development choices, the household discounts future payoffs at exponential rate $\rho = r + \tilde{\alpha}$. This includes both the household's rate of time preference r, and beliefs about the exogenous attrition rate $\tilde{\alpha}$. There is no direct cost of remaining on the waitlist, and no fixed cost of beginning or continuing the application process.[14] The present discounted value to i of being assigned to development j in t years is

$$e^{-\rho t} \, \frac{1}{\rho}(v_{ij} - v_{i0}) \; . \tag{2}$$

4.1.3 Applicant Information and Beliefs

An applicant's optimal initial and final choices will depend on her beliefs about how each possible choice affects the joint distribution of assignments and continued waiting times. Based on institutional features of the Cambridge Mechanism as well as descriptive evidence, I assume that applicants do not know the state of the queue when they first apply, but update their beliefs based on the position information in their final choice letters.[15] When applicant i makes her initial choice, she does so with beliefs about the likely date s and position information p at the final choice stage, whose joint distribution depends on i's initial choice. Let $G_C(s, p)$ denote the probability that the final choice letter is sent less than s years after initial application and that the applicant's list position is no greater than p_j for each development $j \in C$. At the final choice stage, s and p are realized, and i updates her beliefs about the continued waiting time for each development $j \in C$. Let $F_{j,C}(t \mid p)$ denote the probability that continued waiting time for list $j \in C$ is less than t years given position vector p. Importantly, these distributions depend on the full set of initial choice lists C. Due to the algorithm by which the CHA sent out final choice letters, described in Appendix C.1.1, the full set of lists in C could affect the date and information at the final choice stage. In addition, because applicants make their final

[14] Unfortunately, the CHA data and institutional setting did not provide a way to estimate an application cost. This would require exogenous variation in either the value of the application process or the cost of applying.

[15] Appendix B presents evidence that applicants are unaware of short- and medium-term fluctuations in list lengths when they make their initial choices. This finding is also consistent with the information they are given at initial application, and with conversations with the CHA. The CHA generally knew which developments had longer waiting times than others but was unaware of fluctuations in the lengths of particular lists.

choices based on new position information, the full set of list positions p may be informative about the expected continued waiting time for each development $j \in C$.

4.1.4 Choice Problem

Given beliefs and payoffs, an applicant solves the two-stage development choice problem backwards. In the final choice stage, applicant i with initial choice C learns her list positions p and solves

$$\max_{j \in C} \frac{1}{\rho} E\left[e^{-\rho T_j} \mid p\right](v_{ij} - v_{i0}) \tag{3}$$

$$= \max_{j \in C} \int \frac{1}{\rho} e^{-\rho T_j}(v_{ij} - v_{i0}) dF_{j,C}(T_j \mid p) \ . \tag{4}$$

Applicant i makes her initial choice to maximize the expected discounted value of the final choice:

$$\max_{C \in \{0,1,\ldots,J\}^3} E\left[e^{-\rho S} \max_{j \in C} \frac{1}{\rho} E\left[e^{-\rho T_j} \mid P\right](v_{ij} - v_{i0})\right] \tag{5}$$

$$= \max_{C \in \{0,1,\ldots,J\}^3} \int e^{-\rho S} \max_{j \in C} \left[\int \frac{1}{\rho} e^{-\rho T_j}(v_{ij} - v_{i0}) dF_{j,C}(T_j \mid P)\right] dG_C(S, P) \ . \tag{6}$$

Finally, since there is no direct cost of applying or remaining on the waitlist, an eligible household applies for public housing if and only if some development is preferred to her outside option: $\max_j v_{ij} > v_{i0}$. An applicant will also continue the application process if she has not already been removed for exogenous reasons. As a result, counterfactual mechanisms will affect development choices and waiting times, but not which households apply or when they would depart before being offered an apartment.

The empirical strategy requires choosing a specific structure for how applicants make decisions and form beliefs in the Cambridge Mechanism. The development choice model presented here and the belief model in section 5.2 provide one such structure in which applicants are sophisticated and fully account for the complexity of the portfolio choice problem and waitlist dynamics. In absence of direct evidence on how applicants form beliefs and make development choices, it is necessary to make an assumption about this process, and reasonable to use a rational benchmark based on the structure of the decision problem applicants actually faced. However, it is also important to understand to what extent the empirical results are driven by the specific model of decisions and beliefs used in estimation. Appendix C.3 presents results from one alternative model of decisions in which applicants follow a naive heuristic at initial choice. The trade-offs among choice and priority systems are the same as under the sophisticated model used in the main paper.

4.2 Utility Model

The empirical strategy estimates the distribution of flow indirect utilities $v_i = (v_{i1} - v_{i0}, ..., v_{iJ} - v_{i0})$. This section provides a micro-foundation for these payoffs in order to compare changes in utility to the value of cash transfers. The model makes two key assumptions. First, utility is additively separable in housing and non-housing goods, leading to a natural decomposition of flow payoffs into match values and values of assistance. Second, conditional on observed characteristics, applicants whose decisions reveal higher values of assistance have higher marginal utilities of income outside of public housing. In effect, this assumes that the value of assistance is maximally informative about "need." In estimation, I add a restriction on the functional form of utility to parameterize the distribution of v_i and obtain a measure of equivalent variation used for welfare analysis.

4.2.1 Micro-Foundation of Flow Payoffs

Household i receives utility from consumption of housing h and a numeraire c. The utility function is additively separable in the two goods:

$$u(c,h) = u_1(c) + u_2(h)\,.$$

Both u_1 and u_2 are strictly increasing, concave functions. The household has three characteristics: observed income y_i; unobserved income η_i; and development-specific preferences summarized in hedonic indices $h_i = (h_{i1}, ..., h_{iJ})$. Outside of public housing, a household chooses how much to spend on each good given its budget $y_i + \eta_i$. The prices of both goods are normalized to one. The household's flow indirect utility from its outside option is

$$\begin{aligned} v_{i0} &\equiv \max_{c,h} \quad u_1(c) + u_2(h) \quad s.t. \quad c + h \leq y_i + \eta_i && (7) \\ &= v_0(y_i + \eta_i)\,. && (8) \end{aligned}$$

One can think of unobserved income as capturing resources that relax or tighten the household's budget constraint, shifting the value of its outside option. An extensive literature has shown that social ties and alternative living arrangements are an important economic resource for many low-income households (Desmond and An, 2015; Stack, 1974). By modeling these resources as part of the budget constraint, I assume that they are substitutable between housing and the numeraire.

In public housing, household i only has access to observed income y_i. Because it is assigned to a particular apartment, it does not choose how much to spend on housing and the numeraire. Instead, i pays a fixed fraction τ (30 percent) of income in rent, spends the remainder on the numeraire, and enjoys housing consumption h_{ij} in development j. The flow indirect utility from living in development j is

$$v_{ij} \quad \equiv \quad u_1((1-\tau)y_i) + u_2(h_{ij})\,. \tag{9}$$

The difference in flow payoffs is given by

$$v_{ij} - v_{i0} = \underbrace{u_1((1-\tau)y_i) - \overbrace{v_0(y_i + \eta_i)}^{\text{outside option}}}_{\text{value of assistance}} + \underbrace{u_2(h_{ij})}_{\text{match value}}\,. \tag{10}$$

This expression decomposes the difference in flow payoffs into two components: the household's value of assistance and its match value. The value of assistance is common across developments and depends only on household i's observed and unobserved income. It can be thought of as the household's value of the homogeneous aspects of Cambridge public housing. In estimation, the value of assistance will also be allowed to depend on demographic variables such as race/ethnicity and household size. The match value depends on i's taste for the characteristics of development j; it comes from the heterogeneous nature of public housing. These two terms capture the mechanism design trade-off between providing better match quality for housed applicants and housing applicants who want public housing the most. A mechanism that does not give applicants choice over their assignment may induce low-value applicants to reject mismatched offers. If this occurs, more high-value applicants will be housed, with the potential cost that tenants enjoy lower match values.

This utility model embeds two key assumptions. The first is that utility is additively separable in housing and the numeraire. This rules out complementarity between housing and non-housing consumption, and assumes that the match quality a tenant enjoys from their apartment does not affect the value of consuming other goods. The second assumption is that unobserved income is only available outside of public housing, and that it is substitutable between housing and the numeraire. This implies that differences in the value of assistance are driven by households' outside options rather than the value of public housing itself, and that households with the highest values of assistance also have the highest marginal utilities of income (conditional on observables).[16] This assumption has the attractive feature that it maximizes the ability of development choice to reveal a household's level of need. Therefore, the targeting gains from removing choice in Section 7 are an upper bound on what would actually be achieved. Since I find that removing choice can only be justified by a very high value of income redistribution, this finding would be strengthened if choice behavior were less highly correlated with a household's marginal utility of income.

5 Empirical Strategy

This section describes the three steps in estimation. The first step estimates the distribution of potential applicants for Cambridge public housing, including eligible households who did not apply. The second step estimates applicants' beliefs about how their choices affect payoffs through the distribution of assignments and waiting times. The third step estimates preferences over assignments and waiting times by matching application decisions and development choices using the method of simulated moments, taking beliefs and the distribution of potential applicants as inputs (McFadden, 1989; Pakes and Pollard, 1989). A final subsection converts estimates from the utility model to equivalent variation.

[16]One would ideally obtain additional data on applicants' outside options to separate unobserved differences in outside options and tastes for public housing. Such data were not available for this study.

5.1 Distribution of Potential Applicants

The first decision an eligible household makes is whether to apply for public housing at all. Application rates by income and demographic groups will be informative about heterogeneity in the value of assistance. Furthermore, if not everyone applies, public housing applicants will be a selected sample of the eligible population and, in particular, are likely to have worse outside options than observably similar households who did not apply. It is therefore important to account for selection into the applicant pool in preference estimation.

Estimating application rates requires the distribution of characteristics among all households that could have applied for Cambridge public housing during the sample period. This includes households that did apply and also *eligible non-applicants* – eligible households that did not apply and were not already Cambridge public housing applicants or tenants at the beginning of 2010. The CHA dataset contains information on households who applied during the sample period, but not on households which could have applied but did not. Survey data can identify households whose characteristics made them eligible for Cambridge public housing. However, some eligible households were already Cambridge public housing tenants, and others were on the waitlist but applied before 2010. These households were not potential applicants during the sample period, and survey data do not distinguish them from households that could have applied.[17]

My approach is to combine a sample of eligible households from the American Community Survey (ACS) with the CHA dataset to determine the distribution of characteristics among eligible non-applicants. I do this by assigning a probability to each household surveyed in the ACS for whether it appears in the CHA dataset, either as a tenant or as a past or current applicant. These probabilities are a parametric function of household characteristics. I estimate the parameters by matching the characteristics of households in the CHA dataset. One minus each probability is an estimate of the probability that the corresponding ACS household could have applied for Cambridge public housing during the sample period, but did not. Using these probabilities, I draw a sample of eligible non-applicants and combine it with the applicant sample. This procedure is consistent with a model in which households become eligible for public housing once, choose whether to apply, and exit the waitlist or tenancy when they are no longer eligible. Though this model abstracts from the possibility that households might re-apply for public housing, it captures the key idea that households with higher values of living in public housing should be more likely to apply.

The ACS publishes a 5 percent sample of U.S. households covering 2010 through 2014, the same period covered by the CHA applicant dataset.[18] It contains information on household structure and economic and demographic characteristics that determine eligibility and priority for Cambridge public

[17]To my knowledge, no large survey asks households whether they are on a *waitlist* for public housing. The American Community Survey (used here) does ask whether a household receives housing assistance. However, a number of studies including Meyer and Mittag (2015) have shown that these questions tend to understate program participation.

[18]Samples from the ACS can be downloaded here: https://usa.ipums.org/usa-action/variables/group

housing – in particular, whether each ACS household lives or has a member working in Cambridge; whether it meets the income and asset tests; and whether its household structure qualifies it for a two or three bedroom apartment in Family Public Housing.

The probability model parameters are estimated by minimum distance. Households are indexed by $b = 1, ..., B$ and have observed household characteristics Z_b. The ACS assigns each surveyed household a weight w_b based on household b's inverse probability of being sampled – in other words, w_b is the expected number of households that b represents. The estimator chooses a parameter vector θ_{ACS}, which determines the probability that each household appears in the CHA dataset given their characteristics through a probit link function. Denote statistics from the Cambridge dataset by m_{data}, and denote the contribution of each ACS household to the same statistics by m_b. The minimum distance estimator solves

$$\min_{\theta_{ACS}} \ (m_{ACS}(\theta_{ACS}) - m_{data})'(m_{ACS}(\theta_{ACS}) - m_{data})$$

where

$$m_{ACS}(\theta_{ACS}) \equiv \sum_{b=1}^{B} p(Z_b, \theta_{ACS}) w_b m_b$$

$$p(Z, \theta) = \Phi(Z'\theta)$$

5.2 Belief Distributions over Assignments and Waiting Times

The information about preference heterogeneity contained in applicants' development choices depends on their beliefs about how choices affect payoffs. An applicant solving the two-stage development choice problem of Section 4.1 has beliefs about how each initial choice affects the date and position information at the final choice stage, and about continued waiting times for each development given list positions:

$$\{G_C(S, P) \ \ , \ \ \{F_{j,C}(T_j \mid p)\}_{j,p}\}_{C \in \mathcal{C}}$$

Because the final choice stage of the Cambridge Mechanism generates interdependence in waiting times across developments, each possible initial choice may induce a different set of distributions over final choice states and continued waiting times. A major challenge is that data on realized waiting times are sparse, while beliefs are high-dimensional. To address this issue, I assume that applicants have beliefs of a particular form: their beliefs are consistent with the long run steady state distributions that the Cambridge Mechanism would generate given empirical vacancy rates, applicant arrival and departure rates, and initial and final choice frequencies. These empirical quantities can be estimated directly from application data. Combining these estimates with knowledge of the Cambridge Mechanism, I simulate steady state outcomes which quantify interdependence across lists and the option value of the timing and information of the final choice stage. I assume that applicants have these beliefs when simulating the model in the final step of estimation.

The rest of this section describes the model of the Cambridge Mechanism, the construction of simulation inputs, and the construction of belief distributions from simulation outputs.

5.2.1 Structure of Simulation Inputs

Appendix C.1.1 provides a formal model of the Cambridge Mechanism. This section explains the structure placed on inputs that determine assignments. Each day, the following steps occur:

1. New applicants enter the queue and make their initial development choices.
2. Vacant apartments are offered to applicants who have already made their final choices.
3. If the number of applicants on a list who have made their final choices falls below a threshold, the CHA sends final choice letters to a group of applicants on that list. Each letter tells the applicant their current list positions and asks them to make a final choice.
4. Applicants that do not respond to a final choice letter or to an apartment offer are removed from all waitlists.

Given this structure, outcomes in the Cambridge Mechanism are determined by apartment vacancies, arrival and departure dates of applicants, initial and final choices of applicants, and the CHA's policy for sending final choice letters. Vacancies, applicant arrivals and departures, and initial choices do not depend on the state of the waitlist and are modeled as independent exogenous processes; however, the CHA's policy for sending final choice letters and the final choices of applicants do depend on the current state of the waitlist. I therefore place the following structure on inputs:

- Calendar time is indexed in days by $t \in \{1, ..., T\}$. Each list $j \in \{1, ..., J\}$ represents a development and bedroom size. There are S_j apartments represented by list j.
- **Apartment Vacancies:** each vacancy $\nu \in \{1, ..., V\}$ is associated with a calendar date t_ν and a waitlist j_ν. Vacancies occur independently on each list at poisson rates. Vacancy rates were unusually low during the period of study; according to the CHA, the long-run vacancy rate per apartment is once every 10 years, so the vacancy rate of list j is set to $0.1 * S_j$.
- **Applicant Arrivals and Exogenous Departures:** each applicant $i \in \{1, ..., N\}$ arrives on date t_i and becomes unresponsive after date r_i if it has not been housed. Applicants arrive according to a poisson process with arrival rate α. Each applicant becomes unresponsive immediately with probability a_0, and departs at an exponential rate a_1 thereafter.
- **Initial Choices:** applicant i makes an initial choice $C_i \subset \{1, ..., J\}, |C_i| \leq 3$ upon arrival. Since applicants do not know the state of the waitlist when they apply, their initial choices are independent of the current state.
- **Final Choice Letters:** the CHA sends final choice letters according to a rule that depends on the state of each waitlist. For each list j, there is a sequence of trigger and batch size policies

$\{(L_{j,l}, K_{j,l})\}_{l=1}^{L}$ for sending letters. Each day, if fewer than $L_{j,l}$ applicants on list j have made a final choice, this triggers a batch of final choice letters to the next $K_{j,l}$ applicants on list j who have not yet made a final choice. After batch l of final choice letters is sent on list j, pair $(L_{j,l+1}, K_{j,l+1})$ becomes the next trigger and batch policy.

- **Final Choices:** applicants who respond to the final choice letter make their final choice based on their list positions. I use a reduced form model to capture the sensitivity of the final choice to this information. Applicant i selects list $j \in C_i$ with probability

$$\frac{\exp(\beta p_{ij} + \xi_j)}{\sum_{m \in C_i} \exp(\beta p_{im} + \xi_m)}$$

where p_{im} is applicant i's position on list m and ξ_m is a fixed effect for list m.

5.2.2 Construction of Simulation Inputs

The parameters governing inputs are estimated as follows. The annual probability each apartment becomes vacant is calibrated to 10 percent per year.[19] The applicant arrival rate is simply the mean number of applicants per year during the period of study. Initial choice probabilities are also taken directly from the data. Departure parameters were estimated by non-linear least squares using response to the final choice letter as a function of time since application. The coefficients of the final choice model were estimated using the specification in Column (3) of Table 5, replacing continued waiting time with the list position number. Each list has its own distribution of trigger and batch policies, the empirical distribution for the list during the sample period. Sequences of trigger and batch policies are drawn with replacement from their empirical distributions on each list during the period of study.

Given these parameters, I draw sequences of inputs and run the Cambridge Mechanism until it reaches a steady state. Sequences of apartment vacancies and applicant arrival and departure dates are drawn independently. Each applicant's departure date equals its arrival date with probability a_0 and follows an exponential distribution with mean $\frac{1}{a_1}$ years otherwise. The applicant's initial choice is drawn with replacement from the empirical distribution. Finally, I draw a random number for each applicant that determines which final choice it will make given the choice probabilities implied by its list positions.

5.2.3 Construction of Belief Distributions from Simulation Outputs

To construct the relevant distributions from simulation results, I consider what would have happened in the simulation to an additional applicant given each choice the applicant could have made at each stage in the development choice process. For each initial choice, I take the final choice states that would

[19] Due to renovations, the empirical vacancy rate during the sample period was below the long-run average. This approach also assumes an equal vacancy rate per apartment across developments. In principle one could estimate a development-specific vacancy rate based on observed tenant move-outs or the composition of tenants; however, the CHA tenant data do not cover a long enough period for this approach to be effective.

have resulted from that initial choice on a random sample of application dates in the simulation as the distribution $\hat{G}_C(s,p)$. To model the continued waiting time distributions given position information in the final choice stage, $F_{j,C}(T_j \mid p)$, I use a model of continued waiting time that is flexible across initial choices and parametric in list position. For each list j and initial choice C, continued waiting time follows a beta distribution whose parameters depend on current list positions. These distributions are estimated separately for each (j,C) pair using a sample of continued waiting times in the simulation. Appendix C.1 provides details of how these distributions were constructed.

5.3 Preferences over Assignments and Waiting Times

Given the distribution of potential applicants and their beliefs, I estimate the discount factor and parameters governing the distribution of flow payoffs using the method of simulated moments. This section describes the parameterization of flow payoffs, the moments used in estimation, and the construction and minimization of the objective function.

5.3.1 Parameterization of Flow Payoffs

For estimation, I use a Cobb-Douglas utility function:

$$u(c,h) = \gamma \log c + (1-\gamma)\log h\,. \tag{11}$$

Here γ is the fraction of a household's disposable income that it would spend on the numeraire if unconstrained. I also parameterize the distribution of unobserved income η_i and tastes for specific development characteristics h_i. Let Z_i represent observed household characteristics; let X_j represent observed development characteristics; and let X_{ij} represent interactions between applicant and development characteristics. Flow payoffs take the form

$$v_{ij} - v_{i0} = \delta_j + \underbrace{\phi_1 \log y_i - \overbrace{\phi_2 \log(y_i+\eta_i)}^{\text{outside option}} + g(Z_i)}_{\text{value of assistance}} + \underbrace{\sum_k X_{ijk}\beta^o_k + \sum_m X_{jm}\nu_{im}\beta^u_m + \epsilon_{ij}}_{\text{match value}}, \tag{12}$$

where δ_j is a development fixed effect that is common across applicants and (ν_i, ϵ_i) are individual-specific taste parameters not observed by the econometrician. Note that $\phi_1/\phi_2 = \gamma$. The unobserved characteristics are parameterized as

$$\eta_i \overset{iid}{\sim} \begin{cases} TN(0,\sigma^2_\eta, c-y_i, \infty) & w.p.\ \ 1-\Phi\left(\frac{c-y_i}{\sigma_\eta}\right) \\ c-y_i & w.p.\ \ \Phi\left(\frac{c-y_i}{\sigma_\eta}\right) \end{cases} \qquad \nu_{im} \overset{iid}{\sim} N(0,1) \qquad \epsilon_{ij} \overset{iid}{\sim} N(0,1) \tag{13}$$

In addition to placing structure on match values and values of assistance, this parameterization adds development fixed effects and demographic shifters to equation 10. The development fixed effect δ_j captures the component of development quality that is common across households. The value of assistance may depend on other household characteristics Z_i in addition to income. The matching type contains standard terms in discrete choice demand estimation: tastes for observed development characteristics that depend on observed and unobserved household characteristics (v_{im}), and idiosyncratic tastes for each development (ϵ_{ij}).

Unobserved income η_i is parameterized so that observed income is an informative but imperfect predictor of a household's marginal utility of income. With probability $1 - \Phi\left(\frac{\underline{c}-y_i}{\sigma_\eta}\right)$, η_i follows a truncated normal distribution with parameters $(0, \sigma_\eta)$, and with probability $\Phi\left(\frac{\underline{c}-y_i}{\sigma_\eta}\right)$ it is bottom-coded at $\underline{c} - y_i$. The parameter σ_η determines how strongly observed income predicts need: perfectly for $\sigma_\eta = 0$, and not at all as $\sigma_\eta \to \infty$. This parameterization has several attractive features. First, it guarantees each household a minimum consumption level $\underline{c}$ outside of public housing. This is motivated by a variety of social safety net programs that guarantee an individual is not completely destitute. Furthermore, since a positive mass of individuals will be at the consumption minimum, social welfare calculations will not be dominated by a single simulation draw. Second, for each observed income y_i, total income $y_i + \eta_i$ has support on the interval $[\underline{c}, \infty)$. Thus, some low-income households can have low marginal utilities of income while some high-income households can have high marginal utilities. Finally, $E(y_i + \eta_i \mid y_i)$ increases in y_i, an intuitive condition that is consistent with the data.

The parametric restrictions in equation 13 assume independence between values of assistance and match values conditional on observed characteristics, and also place restrictions on the correlation structure of match values across developments. In Section 6.3 I examine the robustness of parameter estimates to more flexible specifications of match value heterogeneity.

5.3.2 Moments and Objective Function

The parameters to be estimated are the discount factor and the parameters governing flow payoffs:

$$\theta \equiv \{\rho, \delta, g(.), \phi, \beta, \sigma_\eta\} .$$

I estimate θ based on moment conditions

$$E[(m_i - E(m_i \mid Z_i, \theta_0)) \mid Z_i] = 0 ,$$

where θ_0 is the true parameter vector, m_i contains features of household decisions, and Z_i contains household characteristics and choice conditions that are determined outside the model. The method of simulated moments captures these conditions in a set of moments, indexed by $q \in \{1, ..., Q\}$, for specific choice features $m_i^{(q)}$ and household characteristics $Z_i^{(q)}$:

$$\hat{g}^{(q)}(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left(m_i^{(q)} - \hat{E}[m_i^{(q)} \mid Z_i, \theta] \right) Z_i^{(q)} .$$

In estimation, the conditional expectation $\hat{E}(m_i \mid Z_i, \theta)$ is estimated by simulation, and the parameter estimate $\hat{\theta}_{MSM}$ is chosen to solve

$$\min_{\theta} \; \hat{\mathbf{g}}(\theta)' \, A \, \hat{\mathbf{g}}(\theta) \, ,$$

where $\hat{\mathbf{g}}(\theta) \equiv (\hat{g}^{(1)}(\theta), ..., \hat{g}^{(Q)}(\theta))'$ and A is a symmetric, positive-definite weight matrix. I match the following choice features $(m_i^{(q)})$ and applicant characteristics $(Z_i^{(q)})$ in the data to those predicted by the simulated model:

1. Application rates by income and demographic groups:

$$m_i^{(q)} = 1\{C_i \neq \emptyset\}; \qquad Z_i^{(q)} = 1\{(y_i, Z_i) \in \mathcal{Y}^{(q)} \times \mathcal{Z}^{(q)}\}$$

2. Development shares among applicants' initial choices: for each list j,

$$m_i^{(q)} = 1\{j \in C_i\}, \qquad Z_i^{(q)} = 1$$

3. Covariances between applicant characteristics and characteristics of their initial development choices:

$$m_i^{(q)} = 1\{C_i \neq \emptyset\} \frac{1}{|C_i|} \sum_{j \in C_i} X_j^{(q)}; \qquad Z_i^{(q)} = 1\{(y_i, Z_i) \in \mathcal{Y}^{(q)} \times \mathcal{Z}^{(q)}\}$$

4. Means and variances of initially chosen development characteristics within and between applicants. An important characteristic is $\bar{T}_j$, the expected waiting time for development j from initial application if an applicant's initial choice was only j.

$$m_i^{(q)} = \frac{1}{|C_i|} \sum_{j \in C_i} X_j^{(q)}, \qquad \left(\frac{1}{|C_i|} \sum_{j \in C_i} X_j^{(q)} \right)^2, \qquad \frac{1}{|C_i|} \sum_{j \in C_i} \left(X_j^{(q)} \right)^2;$$

$$Z_i^{(q)} = 1\{(y_i, Z_i) \in \mathcal{Y}^{(q)} \times \mathcal{Z}^{(q)}\}$$

5. Final choice moments. For all of these, $Z_i^{(q)} = 1$, and $m_i^{(q)}$ depends on the final choice made and the expected continued waiting times of all options in the applicant's choice set $(f_i, \{t_j\}_{j \in C_i})$. I match the fraction of eligible households who made a final choice; the expected continued waiting time of their final choices; the average and maximum difference in expected continued waiting time between the chosen and alternative developments; and a "price index" which measures the waiting time of the final choice relative to other options in the applicant's choice set. The last two sets of moments are intended to capture the response to position information at final choice documented in Section 3.3.

5.3.3 Intuition for Identification

It is useful to consider which moments in the data are most informative about which model parameters. Application rates by income and demographic groups reveal heterogeneity in the value of assistance.

Since low-income and non-white households are more likely to apply for public housing, these groups value living in public housing more on average. However, not all of these households apply, and the rate at which application rates fall with income reveals unobserved differences in values of assistance and/or match values. Initial choices reveal heterogeneity in match values by arguments similar to those in Berry et al. (2004). Covariances between applicant and chosen development characteristics – for example, between an applicant's neighborhood of current residence and the neighborhoods of its chosen developments – reveal which applicants systematically prefer which types of developments. The second moments of chosen development characteristics capture unobserved differences in match values that depend on development characteristics. In addition, the number of and expected waiting times for initially chosen developments reveal unobserved heterogeneity in the value of assistance. Some high income applicants initially choose developments with short waiting times, while others choose long waiting times or select just one or two developments. To the extent that this cannot be explained by observed applicant or development characteristics, or idiosyncratic taste shocks, these differences in behavior suggest that households differ in how much they want public housing overall. Development shares reveal which developments are more desirable conditional on observed characteristics. Finally, combined with the other moments, moments capturing the sensitivity of the final choice to waiting time information separate the discount factor from heterogeneity in flow payoffs.

5.4 Equivalent Cash Transfers

The micro-foundation of preferences provides a way to interpret estimates from the utility model in terms of the value of cash transfers. I use the concept of equivalent variation (EV), the cash transfer that would produce a welfare change equal to that of a public housing assignment, to quantify the welfare gains from alternative allocation policies.

If applicant i is assigned to development j, the equivalent variation EV_{ij} is implicitly defined by

$$v_{ij} - v_{i0} \;=\; v_0(y_i + \eta_i + EV_{ij}) - v_0(y_i + \eta_i)\,, \tag{14}$$

where $v_0(.)$ is the indirect utility function defined in equation 7. Concavity of v_0 implies that an applicant's equivalent cash transfer is increasing in their total income $y_i + \eta_i$, holding the change in flow payoffs $v_{ij} - v_{i0}$ fixed. This is intuitive – higher-income applicants should have greater willingness to pay for the same subjective change in housing quality, for example. Conversely, holding $y_i + \eta_i$ fixed, EV is convex in the change in flow payoffs $v_{ij} - v_{i0}$. As a result, applicants with high flow indirect utility from their assignments require large equivalent transfers.

For the utility function in equation 11, EV has the following closed form expression:

$$EV_{ij} \;=\; (y_i + \eta_i)\left(\exp^{v_{ij} - v_{i0}} - 1\right)\,. \tag{15}$$

It is worth noting that although applicants do not make explicit financial trade-offs in their application and development choices, they do face different *prices* because public housing rent is 30 percent of household income. Differences in application rates by income are therefore informative about households' willingness to pay. Intuitively, if expressed demand for public housing is highly income-dependent, then applicants should have low willingness to pay for specific development characteristics, and EV will primarily depend on the financial subsidy a household receives by living in public housing. In contrast, if income only weakly predicts who applies, this suggests considerable preference heterogeneity relative to the value of the public housing subsidy.[20]

While this reasoning suggests that the conversion from estimated utilities to EV has empirical content, it is not sufficient to obtain an exact conversion. Even in an ideal case where the joint distribution of flow indirect utilities $p(v \mid Z_i)$ was known, the econometrician would require knowledge of applicants' underlying utility functions to identify the marginal utility of income outside of public housing. In the present application, the exact conversion depends on the parametric structure placed on the joint distribution of match values and values of assistance; the Cobb-Douglas functional form of utility; and the assumption that unobserved differences in the value of assistance reflect outside options rather than tastes for public housing itself.

6 Estimation Results

This section presents estimates of the distribution of potential applicants, applicants' waiting time beliefs, and preferences over assignments and waiting times.

6.1 Eligible Population

Appendix Table 14 presents coefficient estimates from the probit model predicting the probability that an ACS household was in the CHA dataset as an applicant or tenant. The probabilities depend on applicant income, race/ethnicity of the household head, and whether the household already lives in Cambridge. The minimum distance estimator matches the total number of households in the CHA dataset; the number of households in six income groups; and the numbers of households from Cambridge and with African American or Hispanic household heads. The point estimates reinforce the discussion of application rates in Section 3.2, with lower-income and non-white households much more likely to appear in the CHA dataset as either applicants or tenants. Though the coefficient estimates have considerable sampling error from the ACS, Figure 6 shows that the pattern of steadily falling application rates by income is consistent across estimates from bootstrapped ACS samples. In addition, the 90 percent confidence interval for the coefficient estimate on an African American household head does not contain zero.

[20]This argument conditions on a particular degree of unobserved heterogeneity in values of assistance (σ_η). This is in part why development choice information is crucial for distinguishing heterogeneity in match values and values of assistance, and why parametric restrictions on the joint distribution of the two objects are needed.

6.2 Applicant Beliefs

Selected parameters governing inputs to the Cambridge Mechanism simulation are shown in Table 13. The annual vacancy rate per unit is calibrated to 10 percent, implying an average of 108 apartment vacancies per year. The applicant arrival rate was 345 per year during the sample period. Based on response to final choice letters, 24.3 percent of applicants become unresponsive immediately, and attrition occurs at a poisson annual rate of 24.5 percent thereafter. Coefficients from the final choice model are also shown. Consistent with the analysis in Section 3.3, applicants are less likely to choose a development with a higher list position.

Table 15 shows the mean and standard deviation of average waiting times for each development in the simulation, and compares them to means in the data. Simulated waiting times are constructed by averaging realized waiting times across applicants housed during the simulation. Simulated waiting times match observed waiting times qualitatively. The largest developments – Jefferson Park, Newtowne Court, Putnam Gardens, and Washington Elms – have simulated average waiting times between 1.0 and 3.2 years. The smaller developments, including Mid and East Cambridge, Lincoln Way, and Jackson Gardens, have longer simulated waiting times of 3.9 to 6.2 years. Although the simulation captures which developments have longer waiting times, the simulated average waiting times are more dispersed than those observed in the data. The main reason for this is that the Cambridge Mechanism was not fully in steady state during the sample period. List closures before and during the sample period allowed some applicants to be housed quickly. In addition, since some developments housed only a few applicants, observed average waiting times have considerable sampling noise. Since applicants had limited information about list closures and current and future fluctuations in list lengths, a reasonable policy would have been to form beliefs based on the long-run distribution of outcomes generated by the Cambridge Mechanism in steady state.

6.3 Preferences over Assignments and Waiting Times

This section presents estimates from three specifications of the development choice model. All specifications include fixed effects for each public housing development, for the race/ethnicity of the household head, and for whether the household currently lives in Cambridge. They include the two terms that depend on income: the value of non-housing consumption while in public housing, and the value of the outside option. They also include an interaction term indicating whether the household lives in the same neighborhood as each development. All specifications include the random effect η_i shifting the value of assistance. Finally, in all specifications and in counterfactuals, households' minimum consumption level $\underline{c}$ is set to \$10 per day. Specification (2) allows for additional observed heterogeneity motivated by differences in application rates and choice patterns. It includes indicators for a three bedroom household, household income below \$20,000, and children below age 10, as well as interactions between development size and Hispanic household head, children under 10, and income below

\$20,000. Specification (3) allows for additional unobserved heterogeneity in match values by adding random coefficients for development size and location. Counterfactuals use estimates obtained from Specification (3). The rest of this section summarizes the parameter estimates and describes features of the preference distribution that will be relevant for counterfactuals.

6.3.1 Parameter Estimates

Applicants are patient. In the first row of Table 6, the estimated annual discount factor is between 0.972 and 0.975 across specifications. The estimates are precise and reject moderate to high degrees of impatience; while applicants exhibit some willingness to substitute towards developments with shorter waiting times, many are willing to wait years for their preferred option. The low discount factor also suggests applicants do not fully anticipate that they might exit the queue before they are housed.

Parameter estimates governing the value of assistance (Panel A of Table 6) show that while income and demographic variables strongly predict the value of public housing, there is substantial unobserved heterogeneity. Consistently across the three specifications, African American headed households have much higher values of living in public housing. Other variables predicting higher values of assistance are intuitive: for example, three bedroom households face higher rents on the private rental market than two bedroom households, while public housing rent depends only on income. Households with a child under age 10 may expect to spend longer in public housing, or to obtain greater benefits for their children.[21] The value of assistance also falls rapidly with observed income: in Specifications (2) and (3), the coefficient estimate on Log of Observed Income implies households would like to spend more than 75 percent of their income on housing. Though large, this estimate is consistent with extremely high rent burdens among low-income households.

Unobserved income makes a substantial contribution to the value of assistance. The point estimates of the scale parameter governing the outside option are greater than \$17,000 in all specifications. In Specification (3), the estimate corresponds to a standard deviation of approximately \$16,052 in the outside options of eligible households, and \$13,716 among applicants.[22] These large estimates are driven by the fact that applicants differ greatly in their selectivity. Some low-income applicants behave as though they can afford to wait a long time for their preferred development, while some high-income applicants appear desperate. These differences cannot be fully explained by observed characteristics or unobserved match value heterogeneity. Due to the high variance in unobserved income, a large fraction (36 percent) of households' outside options are at the consumption minimum $\underline{c}$.

Estimates of match value parameters (Panel B) show substantial heterogeneity in applicants' preferred developments. Location is an important source of predictable heterogeneity, and conditional on other observed characteristics, applicants who already live in Cambridge prefer would actually prefer

[21] Chetty et al. (2015) find that the benefits of living in a neighborhood like Cambridge increase with exposure time.

[22] Recall that σ_η parameterizes a truncated normal distribution to guarantee a minimum consumption level.

to move elsewhere. Hispanic households have a relatively higher value of living in larger developments; other interactions between development size and household characteristics are small in magnitude and imprecisely estimated. Specification (3) estimates a small degree of unobserved heterogeneity in tastes for development size and location. A substantial component of match values are explained by idiosyncratic tastes, with estimated standard deviations of 0.124 to 0.152. Importantly, estimates governing values of assistance – particularly the importance of observed and unobserved income – are stable across specifications of match value heterogeneity.

6.3.2 Features of the Preference Distribution

This section summarizes two features of the preference distribution that drive the trade-off between efficiency and redistribution in counterfactuals: the value of assigning each applicant to their preferred development, and the number of developments for which applicants would accept a take-it-or-leave-it offer. Statistics are based on a sample of applicants drawn from the preference distribution estimated in Specification (3).

There are large welfare gains from matching applicants to their preferred developments. Table 7 displays medians and means of the equivalent variation (EV) from moving an applicant from a lower-ranked choice to their first choice, calculated using equation 15. Across all applicants, the median EV between an applicant's second and first choice is 7.2 percent of observed income, or $563 per year. The mean is even larger, driven by a long right tail in the distribution. These strong preferences for specific developments may be driven by the desire to live in a specific location, for example near a school or workplace, or by other amenities such as building or neighborhood character. The proportional values are similar among African American applicants and applicants with incomes below $15,000, but the dollar values are much lower for low-income applicants. EV from moving an applicant from their last choice to their first choice development is very large, with a median of $1,180 per month across all applicants and $709 per month among low-income applicants.

Because there is substantial heterogeneity in match values, many applicants are only willing to live in a subset of the CHA developments and would reject some take-it-or-leave-it offers of housing. Table 8 tabulates applicants by the number of developments they find acceptable. Some applicants are quite selective – 15 percent would only be willing to live in three or fewer developments – while 50 percent of applicants would be willing to live in any development. Applicants who would accept any development have much lower observed incomes and outside options than more selective applicants. Patterns are qualitatively similar for African American and very low-income applicants, but these groups are less selective overall. 53 percent of African American applicants and 66 percent of applicants with incomes below $15,000 would accept any take-it-or-leave-it-offer.

These statistics suggest that allocation mechanisms which affect match quality and targeting may have large welfare and distributional consequences. A development choice system in which an applicant cannot choose where they live will induce some applicants to reject offers, improving targeting. However, this targeting improvement is likely to come at substantial cost due to lower match quality.

7 Evaluating Design Trade-Offs

Using the estimates from Section 6, I perform counterfactual simulations to evaluate how the development choice and priority systems commonly used to allocate public housing would perform if implemented in Cambridge. These simulations hold all market primitives fixed, including the stock of public housing apartments available to the Cambridge Housing Authority, and only vary allocation policy. I begin with a positive analysis of how these mechanisms would affect total welfare, match quality, and targeting. I then analyze the normative question of which mechanism one should use depending on society's value of income redistribution. This exercise rules out some combinations of choice and priority within a broad class of social welfare functions.

Section 7.1 defines a class of one-stage choice mechanisms that incorporates the range of development choice and priority systems used in practice, and describes the specific mechanisms considered. Section 7.2 presents results from counterfactual simulations of these mechanisms and compares them to a full information benchmark in which the housing authority has full information about applicant preferences.

7.1 Space of Mechanisms

This section formalizes a class of dynamic assignment mechanisms which captures the key features of public housing choice and priority systems used in practice. In this class, applicants make development choices in one stage at initial application, and are ordered on the waitlist lexicographically by priority group and then application date. Compared to the two-stage development choice mechanism used by the CHA, one-stage choice greatly simplifies equilibrium computation and is also more common in practice. To isolate the long-run impacts of policy changes, I abstract from transition dynamics and analyze steady-state equilibria.

The remainder of this section formalizes one stage choice mechanisms, defines equilibrium, explains how allocations are evaluated, and describes the mechanisms explored in counterfactual simulations.

7.1.1 One-Stage Choice Mechanisms

A one-stage choice mechanism φ is defined by two objects:

1. A **development choice system** $\mathcal{C}_\varphi \subseteq 2^{\{1,\dots,J\}}$. Each element of $\mathcal{C}_\varphi$ is a subset of developments from which the applicant may receive apartment offers.
2. A **priority system** $\psi_\varphi : \mathcal{Z} \longrightarrow \{1, ..., B\}$ maps applicant characteristics to a priority group. Applicant i has higher priority than applicant i' in φ if $\psi_\varphi(Z_i) < \psi_\varphi(Z_{i'})$.

The mechanism operates on sequences of apartment vacancies, applicant arrivals, and exogenous applicant departures. Each vacancy $\nu \in \{1, ..., V\}$ has a date t_ν and development j_ν. Each applicant $i \in \{1, ..., N\}$ has arrival date t_i, departure date r_i, observed characteristics Z_i, and payoff vector $v_i = (v_{i0}, v_{i1}, ..., v_{ij})$. The mechanism φ runs according to the following algorithm. On each date t,

(i) Each arriving applicant ($t_i = t$) chooses a set of developments $C_i \in \mathcal{C}_\varphi$ and is placed on the waitlist for each development $j \in C_i$. On each list, applicants are ordered lexicographically by $(\psi_\varphi(Z_i), t_i)$.

(ii) Each vacancy ν with $t_\nu = t$ is offered to the first applicant on list j_ν. If the applicant accepts, it is housed and removed from all lists $j \in C_i$. If the applicant rejects, it is removed from all waitlists and cannot reapply. This step is repeated until an applicant accepts or the waitlist is empty. If the latter occurs, the vacancy is held until the next day.

(iii) Departing applicants ($r_i = t$) are removed from all lists $j \in C_i$.

7.1.2 Development Choice Problem, Information, and Equilibrium

In a one stage choice mechanism, an applicant simply considers, for each possible subset of developments they can choose, which development is likely to arrive first and the distribution of waiting times for the first arrival. Let T_j be the random variable for the waiting time for development j if an applicant were only on the waitlist for j. The realization of T_j will depend on applicant i's date of application. The joint distribution $F_{T_1,...,T_J}$ may depend on the applicant's priority $\psi_\varphi(Z_i)$. The applicant solves the following choice problem:

$$\max_{C \in \mathcal{C}_\varphi} \sum_{j \in C} w_j^C(\psi_\varphi(Z_i))(v_{ij} - v_{i0}) \tag{16}$$

$$w_j^C(\psi_\varphi(Z_i)) \equiv \frac{1}{\rho} E_{\psi_\varphi(Z_i)}\left[e^{-\rho T_j} \mid T_j = \min_{k \in C_i} T_k\right] \times P_{\psi_\varphi(Z_i)}\left[T_j = \min_{k \in C_i} T_k\right] \ . \tag{17}$$

As in the Cambridge Mechanism, I assume applicants do not know the state of the queue when they apply, and instead have a common prior over the distribution of outcomes that they face for each possible choice $C \in \mathcal{C}_\varphi$ given their priority group $\psi_\varphi(Z_i)$ and the mechanism's steady state. As a result, beliefs do not depend on application date. In equilibrium, beliefs are consistent with the distributions generated by the mechanism in long-run steady state given the distribution of potential applicants, the preference distribution $p(v_i \mid Z_i, \hat{\theta}_{MSM})$, and given that applicants choose developments according to equation 16.

In the counterfactual simulations, the exogenous departures model is the same as in the Cambridge Mechanism simulation, as are vacancy rates. Applicant arrivals are generated using the distribution of potential applicants and preferences estimated in Section 6. Each applicant's choice solves equation 16 given preferences and beliefs. Potential applicants choose to apply if any development is preferable to their outside option. Appendix D provides details of how the equilibrium is computed. The algorithm iteratively updates applicant choices and their implied steady state waiting time distributions until a fixed point is reached.

7.1.3 Evaluating Allocations

Given sequences of inputs, a mechanism φ produces an eventual assignment $j_\varphi(i) \in \{0, 1, ..., J\}$ for each applicant, with $j_\varphi(i) = 0$ if applicant i is not assigned an apartment. A natural way to summarize the welfare and distributional impacts of a mechanism is to average characteristics of assigned applicants and their values over assigned apartments. In long-run steady state, if applicants vacate apartments at a common exogenous, poisson rate, then this provides an estimate of the mean characteristics of public housing tenants at any given time. A social planner interested in maximizing the expected discounted sum of future payoffs would be interested in these statistics. To summarize welfare, I use equivalent cash transfers as a baseline measure:

$$W(\varphi) = \frac{1}{\sum_{i=1}^{N} 1\{j_\varphi(i) \neq 0\}} \sum_{i=1}^{N} EV_{i,j_\varphi(i)} \tag{18}$$

where $EV_{i,j_\varphi(i)}$ is as defined in equation 15. To summarize characteristics of housed applicants, one can do the same for transformations of applicant characteristics:

$$\frac{1}{\sum_{i=1}^{N} 1\{j_\varphi(i) \neq 0\}} \sum_{i=1}^{N} h(Z_i, v_i, j_\varphi(i)) \tag{19}$$

To incorporate social welfare weights into welfare calculations, one can transform equivalent variation from assignments by a function $f(Z_i, v_i, EV)$ that depends on applicant characteristics:

$$W(\varphi; f) = \frac{1}{\sum_{i=1}^{N} 1\{j_\varphi(i) \neq 0\}} \sum_{i=1}^{N} f(Z_i, v_i, EV_{i,j_\varphi(i)}) \tag{20}$$

In particular, this formulation allows a social planner to have different marginal values of transferring one dollar to different households.

Finally, one can compare welfare gains from different mechanisms adjusting for the total cost of the public housing program. Since rent in public housing is proportional to a tenant's income, the CHA will receive lower rent payments if it houses lower-income applicants. Estimating the fiscal cost of public housing is challenging for a variety of reasons.[23] Instead, I use market rents in Cambridge, MA during the sample period as a proxy for the fiscal opportunity cost of the public housing program. Between 2010 and 2014, a conservative estimate of the market rent for a modest two- to three-bedroom

[23] Olsen (2009) provides a thorough discussion of the complexity of estimating the fiscal cost of providing public housing. One of the primary challenges is that public housing is a durable good which depreciates over time and requires lumpy investment to maintain. Annual maintenance and administrative costs will dramatically under- or over-state the true cost of public housing, depending on the year they are measured. In addition, they will not capture opportunity costs from revenue that might be generated by other uses of the land and buildings owned by a PHA. For example, selling Cambridge Housing Authority land to private developers would generate not only revenue from the sale, but also property tax revenue for the city of Cambridge.

apartment was \$2,000 per month, or $c \equiv \$24,000$.[24] Subtracting tenant rent payments from this cost measure provides a reasonable lower bound. Adjusted for cost, welfare gains are

$$\tilde{W}(\varphi; f) = \frac{\sum_{i=1}^{N} f(Z_i, v_i, EV_{i,j_\varphi(i)})}{\sum_{i=1}^{N} 1\{j_\varphi(i) \neq 0\}(c - 0.3y_i)} \tag{21}$$

7.1.4 Simulated Mechanisms

The mechanisms used by the 24 surveyed PHAs in Section 2 can be modeled using six development choice systems and three priority systems. I computed the counterfactual equilibrium that would arise in Cambridge under each combination. The development choice systems are

1. **Choose One**: $\mathcal{C} = \{\{1\}, ..., \{J\}\}$. Applicants must select one development. This choice system is closest to those used in Cambridge, New York City, New Haven, and Seattle, which allow applicants to select a limited number of developments.
2. **Choose Any Subset**: $\mathcal{C} = 2^{\{1,...,J\}}$. Applicants may choose any subset of developments, as in Boston and San Antonio.
3. **Choose All or One**: $\mathcal{C} = \{\{1\}, ..., \{J\}, \{1, ..., J\}\}$. Applicants may either wait for their preferred development or take the first available offer from any development. This choice system approximates the policies used in Philadelphia, Baltimore, and Newark.
4. **Choose Neighborhood**: $\mathcal{C} = \{C_{\text{north}}, C_{\text{east}}, C_{\text{central}}\}$. Applicants choose a neighborhood from which to receive an apartment offer. Importantly, an applicant cannot choose to wait for their most preferred development.
5. **Choose All or Neighborhood**: $\mathcal{C} = \{C_{\text{north}}, C_{\text{east}}, C_{\text{central}}, \{1, ..., J\}\}$. Applicants may either choose a neighborhood or receive the first offer city-wide. Chicago uses this development choice system for Family Public Housing.
6. **No Choice**: $\mathcal{C} = \{\{1, ..., J\}\}$. Applicants must accept the first available apartment in any development; they have no choice over their assignment.

For priority systems, I model priority for higher socioeconomic status households as a priority for higher-income applicants, and lower socioeconomic status or need-based priorities as a priority for lower-income applicants:

1. **Equal Priority**: Applicants are treated equally and ordered only by application date. Apart from emergency priorities that affect few applicants, several PHAs, including the CHA, use equal priority.
2. **Low-Income Priority**: Applicants with household incomes below \$15,000 are offered apartments first. Seattle prioritizes households below 30 percent AMI, and several PHAs use "need-based"

[24] Based on the Zillow Rent Index: `https://www.zillow.com/cambridge-ma/home-values/`. Median rents ranged from about \$2,200 to \$2,600 for two-bedroom units, and \$2,600 to \$2,900 for three-bedroom units.

priorities for households that are severely rent burdened, face involuntary displacement, or are referred by other agencies that provide public assistance.

3. **High-Income Priority**: Applicants with household incomes above $15,000 are offered apartments first. New York City and New Haven explicitly prioritize households with higher incomes. High-Income Priority also captures priorities for working or economically self-sufficient households used by several PHAs.

7.2 Welfare and Distributional Impacts of Allocation Policy

I begin by analyzing the effect of development choice systems under Equal Priority and then consider the effects of prioritizing higher- or lower-income applicants. Finally, I show how distributional preferences determine which mechanism should be adopted in Cambridge.

7.2.1 Effect of Development Choice under Equal Priority

The range of development choice systems used in practice involves a dramatic trade-off between match quality and targeting. Columns (1) and (6) of Table 9 summarize allocations under Choose One and No Choice with Equal Priority. Appendix tables 19 and 20 present results for the same development choice systems under Low- and High-Income Priority. Under Choose One, the average tenant values their assignment as much as an annual cash transfer of $15,243, or $1,270 per month; under No Choice, the value falls by 32 percent to $10,406. Part of this welfare loss is driven by a reduction in match quality. While 52 percent of tenants are assigned to their first choice development under Choose One, only 8 percent are under No Choice. However, No Choice substantially improves targeting by inducing applicants with higher incomes and better outside options to reject offers. The mean observed income of tenants falls from $18,252 to $16,903, and outside options fall by even more. Due to lower tenant incomes, the CHA would receive lower rent payments and therefore incur a higher cost per unit under No Choice. Adjusted for cost, Choose One achieves 82 cents of welfare gains per dollar spent, while No Choice achieves 55 cents, a 33 percent decrease.

Other development choice systems produce intermediate allocations in terms of match quality and total welfare. Choose Any Subset and Choose All or One, which allow applicants to select several developments as a hedge against waiting time uncertainty, have modest effects on the allocation. This is because in equilibrium, waiting time uncertainty is small relative to differences in average waiting times across developments. Applicants that choose several developments are very likely to be housed in the development with the shortest expected waiting time, and would have picked that development under Choose One. In contrast, Choose Neighborhood and Choose All or Neighborhood, which allow applicants to choose their neighborhood but not a specific development, do impact assignments. Each neighborhood contains at least three developments, so some applicants reject offers, lowering match quality and improving targeting relative to Choose One.

Table 9 illustrates an important exception to the trade-off between match quality and targeting broadly reflected in these results. The fraction of extremely high-need tenants, who would be at the minimum consumption level outside of public housing, is maximized under Choose Neighborhood (50 percent) rather than under No Choice (47 percent). This occurs because choice provides an additional way for applicants to increase their chances of receiving an assignment: by selecting a less demanded neighborhood, an applicant lowers their expected waiting time and hence the probability of exit. This option is not available under No Choice because all applicants face the same waiting times. Under Equal Priority, the targeting gain through differential waiting times under Choose Neighborhood outweighs the gain from additional rejections under No Choice.[25] This finding illustrates that there is not always trade-off between match quality and targeting. When alternatives differ in their average desirability, allowing choice may improve both targeting and match quality.[26]

7.2.2 Effect of Income-Based Priorities

Income-based priorities strongly affect targeting but only modestly influence match quality and efficiency. Columns (1) - (6) of Table 10 summarize allocations under Low-Income, High-Income, and Equal Priority for the Choose One and No Choice development choice systems. These mechanisms illustrate the range of outcomes achieved under all mechanisms considered in the paper.

The priority system dramatically affects tenants' observed incomes and outside options. Under Choose One, High-Income Priority leads to an average tenant income of \$22,306, compared to \$13,505 under Low-Income Priority. The fraction of extremely high-need tenants also rises from 40 percent to 51 percent. Because observed income is a strong predictor of need, the mechanism designer can achieve greater targeting gains by using observed characteristics through the priority system than by inducing self-selection through the development choice system. Choice and priority work together to maximize targeting: under Low-Income Priority, No Choice, 57 percent of tenants extremely high-need.

Compared to development choice, the effects of income-based priorities on match quality and overall efficiency are more modest. Equivalent variation per assigned unit is maximized under Low-Income Priority, Choose One (\$15,444). This is intuitive since the lowest-income applicants receive the largest rent subsidies as tenants. Due to the differential cost of housing lower-income tenants, EV per dollar spent is highest under Equal Priority, Choose One (82 cents per dollar). However, the differences across priority systems are small: EV per dollar ranges from 75 to 82 cents under Choose One and 53 cents to 57 cents under No Choice.

Priorities do affect match quality by differentially shifting incentives for higher- and lower-priority applicants. Under Choose One, the fraction of tenants assigned to their first choice development is

[25]This finding depends on the priority system. Appendix Table 19 shows that under Low-Income Priority, No Choice maximizes the fraction of extremely high-need tenants.

[26]Reeling and Verdier (2018) find a stronger version of this result in the context of hunting licenses, where sites are highly vertically differentiated. The estimates obtained here find that CHA developments are mostly horizontally differentiated, and as a result, the dominant trade-off is between match quality and targeting.

42 percent under Low-Income Priority, 45 percent under High-Income Priority, and 52 percent under Equal Priority. The moderate reduction in match quality under High- and Low-Income Priority relative to Equal Priority reflects two countervailing forces: high-priority individuals become more selective while low-priority individuals become less so. The overall effect of introducing priorities on match quality depends both on how the proportions of high- and low-priority tenants change, and on how the development choices of each group are affected.

7.2.3 Incorporating a Preference for Redistribution

Given that alternative choice and priority systems can yield dramatically different allocations, how should a PHA decide which mechanism to use? Measuring welfare gains using equivalent variation places equal value on transferring resources to households at different points in the income distribution. Since public housing is an anti-poverty program, it is natural to allow a PHA to value income redistribution, i.e. to prefer making transfers to households with higher marginal utilities of income.[27] This section explicitly incorporates distributional preferences into welfare comparisons among allocation mechanisms.

In the preference model presented in Section 4.2, a mechanism designer who wishes to transfer to households with higher marginal utilities of income should apply higher social welfare weights to households with worse outside options. A household's utility from its outside option is determined by its total income $\tilde{y}_i \equiv y_i + \eta_i$. Any monotonically increasing function $f(\tilde{y}_i)$ corresponds to a social welfare function which penalizes inequality. To capture a wide range of distributional preferences, I consider a parametric class of social welfare functions proposed by Atkinson (1970):

$$f(\tilde{y}_i, EV; \lambda) = \begin{cases} \dfrac{1}{1-\lambda}\left[(\tilde{y}_i + EV)^{1-\lambda} - \tilde{y}_i^{1-\lambda}\right] & \lambda \neq 1 \\ \log(\tilde{y}_i + EV) - \log(\tilde{y}_i) & \lambda = 1 \end{cases}$$

This class of functions exhibits "constant relative inequality aversion," with the degree of inequality aversion parameterized by the scalar λ. The social welfare generated by transferring one dollar to a household with 1 percent lower income is approximately λ percent greater. An inequality aversion parameter of $\lambda = 0$ implies no taste for redistribution; $\lambda = \infty$ corresponds to a social welfare function that only values welfare changes for the worst-off agents. For $\lambda > 0$, social welfare increases whenever resources are transferred from higher- to lower-income households, and for any $\lambda \in \mathbb{R}$ income distributions are ranked identically if incomes are multiplied by a constant. For a given value of λ, one can use equation 21 to rank mechanisms according to the social welfare generated by the resulting allocations.

Figure 1 shows that under the current CHA priority system (Equal Priority), applicants should have some choice over where they live for any $\lambda > 0$. The figure plots the cost-adjusted welfare gains

[27] In the model presented in Section 4.2, the most efficient way to achieve income redistribution would be through the income tax schedule (Atkinson and Stiglitz (1976)). The present analysis assumes that the tax schedule is not available to the policy maker as an instrument; rather, the policy maker can only increase social welfare through the allocation of its existing public housing stock.

from equation 21 for each mechanism, normalized by welfare under Equal Priority, Choose One at each value of λ. Consistent with Table 9, Choose One is preferred with low inequality aversion because it produces the highest EV per dollar spent. With high inequality aversion, Choose Neighborhood is preferred because it maximizes the proportion of extremely high-need tenants. Under Equal Priority, No Choice is suboptimal for any degree of inequality aversion. Appendix Figure 7 shows that under Low-Income Priority, No Choice is preferred to Choose Neighborhood with high inequality aversion.

Figure 2 compares alternative priority systems under Choose One. With a low degree of inequality aversion, Equal Priority is preferred because it maximizes EV per dollar spent. As inequality aversion increases, the CHA should prioritize low-income applicants because the social value of welfare gains to lower-income households outweighs both the additional cost of housing them and the moderate reduction in match quality.

To determine which mechanism would be preferred under different social welfare functions, Figure 3 plots the mechanisms which perform best for some degree of inequality aversion. With low inequality aversion, the CHA should prioritize households equally and allow choice. As λ increases, the CHA should first prioritize low-income applicants and still allow choice, and then, if its taste for redistribution is sufficiently high, begin restricting applicants' ability to choose their preferred developments. With a very high value of λ, Low-Income Priority, No Choice performs best. The targeting gains of this mechanism become worth the match quality distortion from not allowing choice as well as the additional cost of housing lower-income tenants.

Although the preferred mechanism depends on distributional preferences, several combinations of choice and priority are strictly dominated in the Cambridge setting: for any degree of inequality aversion, there is a better policy. Figure 4 plots the upper envelope from Figure 3 along with a subset of dominated mechanisms. Two mechanisms which perform particularly poorly are No Choice with Equal and High-Income Priority. Intuitively, these mechanisms produce the worst of both worlds. No Choice creates an enormous match quality distortion, decreasing tenant welfare and increasing the fiscal cost of the program. If the social planner values redistribution enough to justify No Choice, it should also prioritize Low-Income applicants, since doing so improves targeting with a much smaller reduction in match quality. A mechanism such as the one used in Los Angeles, which combines No Choice with priority for economically self-sufficient households, is strictly sub-optimal in Cambridge within this class of social welfare functions. In fact, Appendix Figure 11 shows that every mechanism with High-Income Priority is dominated in Cambridge. The lower cost of housing higher-income tenants does not offset their lower values of the assistance they would receive.

Finally, the Cambridge Mechanism is likely to perform well under low to moderate inequality aversion. Apart from the two stages of choice, the Cambridge Mechanism has the same structure as Equal Priority, Choose One. If the allocations produced by the two mechanisms are similar, then the Cambridge Mechanism nearly optimal among the mechanisms considered at inequality aversion parameters below 1.2. This range includes values of λ implying a reasonably strong taste for redistribution: when $\lambda = 1.2$, the social value of transferring \$2.30 to a household earning \$20,000 per year equals the value of transferring just \$1 to a household earning \$10,000 per year.

7.2.4 Full-Information Benchmark

Another important question is how well the CHA could do if it obtained more information about applicants. Columns (7) and (8) of Table 9 provide a lower bound on the welfare and targeting gains that would be possible if the social planner fully knew current applicants' preferences and outside options, but did not know when applicants would arrive and depart in the future. The results show that private information sharply limits what can be achieved. The social planner maximizes equivalent variation per assigned unit in Column (7) and minimizes the outside options of tenants in Column (8). In both cases, the planner uses a greedy algorithm, housing the applicant with the highest social value when an apartment becomes available without taking dynamic considerations into account.

In the EV-maximizing allocation, assignments are valued 17 percent more highly than under Low-Income Priority, Choose One. The solution focuses on housing applicants with the highest values of assistance rather than on maximizing match quality. In fact, the EV-maximizing allocation achieves lower match quality than the Choose One mechanism, but better targeting than *any* mechanism considered here. 23.5 percent of tenants live in their first choice development under EV-maximization, about half the proportion under Choose One. Meanwhile, 61.7 percent of tenants are Extremely High-Need, 5 percent more than under Low-Income Priority, No Choice. Tenants are more likely to have an African American household head and be from Cambridge, which predict a high value of assistance. The targeting-maximizing allocation sacrifices match quality completely, but still achieves moderate welfare gains. 80 percent of tenants are Extremely High-Need, but only 7 percent live in their first choice development. EV per unit and per dollar spent are lower than under the Choose One mechanisms, but greater than under No Choice.

These results suggest that the potential gains from using more detailed information about applicants in allocation could be substantial. In fact, many PHAs already use need-based priorities that are strongly predictive of need. For example, PHAs prioritize victims of domestic violence, the homeless, or households that are severely rent burdened or have been involuntarily displaced. In principle, PHAs could obtain additional information about applicants to better predict outside options or preferred developments. However, detailed priorities also raise the possibility that households might manipulate their characteristics to obtain higher priority. This important concern is left for future work.

8 Conclusion

The allocation of scarce public resources involves trading off efficiency and other policy goals, such as fairness or redistribution. This paper empirically studies a trade-off arising in the allocation of public housing between efficiency and income redistribution. Using data on the choices of public housing applicants in Cambridge, MA, I estimate a model of preferences for public housing that quantifies heterogeneity in applicants' preferred developments and overall values of obtaining assistance. The empirical strategy exploits a trade-off faced by applicants between shorter waiting times and preferred assignments as well as the structure of the allocation mechanism used in Cambridge. I use the estimated model to simulate counterfactual equilibria under allocation mechanisms in use across the U.S.

In Cambridge, the range of choice and priority systems used in practice would dramatically affect efficiency and redistribution. Mechanisms allowing applicants to choose their preferred development can provide large welfare gains to tenants, comparable to cash transfers of more than $1,200 per month. Mechanisms that do not allow choice would induce many applicants to reject mismatched apartment offers, allowing more disadvantaged applicants to be housed. This would lower match quality for tenants, and cost-adjusted welfare gains would fall by more than 30 percent. The CHA could achieve the same increase in targeting without lowering tenant welfare by prioritizing lower-income applicants and allowing choice. As a result, some of the mechanisms used in other cities are strictly dominated in Cambridge within a broad class of social welfare functions. Prioritizing higher-income applicants without allowing choice, as is done in some cities, is suboptimal whether society has a high or a low value of income redistribution.

These findings yield concrete policy takeaways for housing authorities. A number of papers have argued that ordeals can increase the efficiency of public programs by more effectively targeting intended beneficiaries. My results weigh against PHAs using development choice restrictions as an ordeal. Because choice restrictions impose a large cost on tenants, a policy maker should only use them with very strong preferences for redistribution. In addition, PHAs already collect applicant information that is highly predictive of need, and they can use this information to improve targeting without creating inefficient matches for tenants. PHAs should only use choice restrictions after establishing priorities based on these observed characteristics.

This study also raises a number of questions for future work on the design of dynamic allocation mechanisms and government-provided benefits. Optimal dynamic mechanisms in settings like public housing allocation are an open theoretical question, especially when applicants can manipulate their characteristics in response to priorities. Theoretical insights into optimal mechanisms could provide policy guidance for PHAs and other organizations which allocate scarce resources over time. A complementary direction for future work is to ask how housing assistance benefits should themselves be designed. Would it be better to provide less generous public housing benefits but cover more eligible households? Should housing assistance be provided in-kind through public housing, or through private market subsidies as in the Housing Choice Voucher program? Should the government provide housing-specific subsidies at all? The revealed preference methods developed here may prove useful for answering such questions.

References

Abdulkadirolu, Atila, Nikhil Agarwal, and Parag A. Pathak, "The Welfare Effects of Coordinated Assignment: The New York City High School Match," *American Economic Review*, 2017, *95* (2), 364–367.

Agarwal, Nikhil, "An Empirical Model of the Medical Match," *American Economic Review*, 2015, *105* (7), 1939–1978.

_ **and Paulo Somaini**, "Demand Analysis using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, 2018, *86* (2), 391–444.

_ **, Itai Ashlagi, Michael Rees, Paulo Somaini, and Daniel Waldinger**, "An Empirical Framework for Sequential Assignment: the Allocation of Deceased Donor Kidneys," *NBER Working Paper 25607*, 2019.

Akerlof, George, "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *American Economic Review*, 1978, *68* (1), 8–19.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi, "Self-Targeting : Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, 2016, *124* (2).

Andersson, Fredrik, Mark J. Kutzbach, Giordano E. Palloni, Henry O. Pollakowski, and Daniel H. Weinberg, "Childhood Housing and Adult Earnings: A Between-Siblings Analysis of Housing Vouchers and Public Housing," *NBER Working Paper 22721*, 2016.

Arnosti, Nick and Peng Shi, "Design of Lotteries and Waitlists for Affordable Housing Allocation," 2018.

Atkinson, Anthony B, "On the Measurement of Inequality," *Journal of Economic Theory*, 1970, *2*, 244–263.

Atkinson, Anthony B. and Joseph E. Stiglitz, "The Design of Tax Structure: Direct Versus Indirect Taxation," *Journal of Public Economics*, 1976, *6*, 55–75.

Berry, Steven, Ariel Pakes, and James Levinsohn, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 2004, *112* (1), 68–105.

Bloch, Francis and David Cantala, "Dynamic Assignment of Objects to Queuing Agents," *American Economic Journal: Microeconomics*, 2017, *9* (1), 88–122.

Chade, Hector and Lones Smith, "Simultaneous Search," *Econometrica*, 2006, *74* (5), 1293–1307.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2015, *106* (4), 1–88.

Collinson, Robert and Davin Reed, "The Effects of Residential Evictions on Low-Income Adults," 2018.

__ **, Ingrid Gould Ellen, and Jens Ludwig**, "Low-Income Housing Policy," *NBER Working Paper 21071*, 2015.

Deshpande, Manasi and Yue Li, "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Policy*, 2019, pp. 1–39.

Desmond, Matthew, "Disposable Ties and the Urban Poor," *American Journal of Sociology*, 2012, *117* (5), 1295–1335.

__ **and Weihua An**, "Neighborhood and Network Disadvantage among Urban Renters," *Sociological Science*, 2015, *2*, 329–350.

Dworczak, Piotr, Scott Duke Kominers, and Mohammad Akbarpour, "Redistribution through Markets," 2018.

Fack, Gabrielle, Julien Grenet, and Yinghua He, "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions," *American Economic Review*, 2019.

Finkelstein, Amy and Matthew J Notowidigdo, "Take-up and Targeting: Experimental Evidence from SNAP," 2018.

Geyer, Judy and Holger Sieg, "Estimating a Model of Excess Demand for Public Housing," *Quantitative Economics*, 2013, *4* (3), 483–513.

Glaeser, Edward L. and Erzo F.P. Luttmer, "The Misallocation of Housing Under Rent Control," *American Economic Review*, 2003, *93* (4), 1027–1046.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger, "Heterogeneous Preferences and the Efficacy of Public School Choice," 2009.

He, Yinghua, "Gaming the Boston School Choice Mechanism in Beijing," 2017.

Humphries, John Eric, Nick Mader, Daniel Tannenbaum, and Winnie Van Dijk, "Does Eviction Cause Poverty ? Quasi-Experimental Evidence from Cook County, IL," 2018.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Leiber, Ethan and Lee Lockwood, "Targeting with In-Kind Transfers: Evidence from Medicaid Home Care," *The American Economic Review*, 2019.

Leshno, Jacob D, "Dynamic Matching in Overloaded Waiting Lists," 2017.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu, "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity," *American Economic Review, Papers & Proceedings*, 2013, *103* (3), 226–231.

McFadden, Daniel, "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 1989, *57* (5), 995.

Meyer, Bruce and Nikolas Mittag, "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness and Holes in the Safety Net," *Upjohn Institute Working Paper 15-242*, 2015.

Narita, Yusuke, "Match or Mismatch? Learning and Inertia in School Choice," 2018.

Nichols, Albert L. and Richard J. Zeckhauser, "Targeting Transfers through Restrictions on Recipients," *The American Economic Review Papers and Proceedings*, 1982, *72* (2), 372–377.

Olsen, Edgar O., "The Cost-Effectiveness of Alternative Methods of Delivering Housing Subsidies," 2009.

Pakes, Ariel and David Pollard, "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 1989, *57* (5), 1027–1057.

Reeling, Carson and Valentin Verdier, "Welfare Effects of Dynamic Matching: An Empirical Analysis," 2018.

Sieg, Holger and Chamna Yoon, "Waiting for Affordable Housing in New York," 2016.

Stack, Carol, *All Our Kin: Strategies for Survival in a Black Community.*, New York: Basic Books, 1974.

Thakral, Neil, "The Public-Housing Allocation Problem : Theory and Evidence from Pittsburgh," 2016, (August), 1–55.

Van Dijk, Winnie, "The Socio-Economic Consequences of Housing Assistance," 2018.

Van Ommeren, Jos N. and Arno J. Van der Vlist, "Households' Willingness to Pay for Public Housing," *Journal of Urban Economics*, 2016, *92*, 91–105.

Table 1: Allocation Policies Used by Public Housing Authorities

Public Housing Authority (PHA) Jurisdiction	City Population in 2016	Number of Public Housing Units in 2013	Priority System	Development Choice System
Panel A: PHA's with Largest Public Housing Stock				
New York City, NY	8,537,673	175,000	Mixed	Limited Choice
Chicago, IL	2,704,958	21,150	Equal	Limited or All
Philadelphia, PA	1,567,872	15,000	Equal	Limited or All
Baltimore, MD	614,664	11,250	High SES	Limited or All
Boston, MA	673,184	10,250	Equal	Any Subset
Cleveland, OH (Cuyahoga Metro Area)	385,809	10,000	High SES	Limited Choice
Miami, FL	453,579	9,400	Equal	No Choice
Washington, D.C. *	681,170	8,350	--	--
Newark, NJ	281,764	7,750	High SES	Limited or All
Los Angeles, CA	3,976,322	6,900	High SES	No Choice
Seattle, WA	704,352	6,300	Low SES	Limited Choice
Minneapolis, MN	413,651	6,250	Low SES	No Choice
San Antonio, TX	1,492,510	6,200	Low SES	Any Subset
Panel B: PHA's comparable to Cambridge, MA (2000-3000 public housing units, 100-200K population)				
Cambridge, MA	110,650	2,450	Equal	Limited Choice
Rochester, NY *	114,011	2,500	Equal	No Choice
New Haven, CT	129,934	2,600	High SES	Limited Choice
Columbia, SC	134,209	2,140	Equal	No Choice
Dayton, OH	140,489	2,750	High SES	Any Subset
Syracuse, NY *	143,378	2,340	High SES	No Choice
Bridgeport, CT *	145,936	2,600	Equal	--
Kansas City, KS	151,709	2,050	Mixed	No Choice
Macon, GA *	152,555	2,250	High SES	No Choice
Providence, RI	179,219	2,600	Equal	No Choice
Worcester, MA *	184,508	2,470	Low SES	No Choice
Augusta, GA *	197,081	2,250	Equal	No Choice
Yonkers, NY	200,807	2,080	Equal	Any Subset

Notes: Features of allocation mechanisms used by PHAs in 25 cities. PHAs were chosen based on city population and/or the size of their public housing stocks. * indicates that the PHA's administrative plan was not available online. In these cases, information was gleaned from the PHA website and application forms. A High SES priority system favors households above 30% of Area Median Income (AMI), or which are economically self-sufficient or have a working member. A Low SES priority system prioritizes households below 30% AMI, or which are severely rent burdened or have been involuntarily displaced. A Mixed priority system prioritizes some (but not all) households of both types, and an Equal priority system does not prioritize households based on socioeconomic status. Under Limited Choice, applicants must choose a small number of developments from which to receive offers. Under Any Subset, applicants may choose any subset of the developments. Under No Choice, applicants must accept the first available apartment in any development. Under Limited or All, applicants may either commit to taking the first available apartment or select a limited number of developments. In Chicago, applicants for Family Public Housing may select a specific neighborhood, but not developments within a neighborhood.

Table 2: Characteristics of Eligible and Applicant Households

	All		by Year of Initial Application				
	Eligible	Applied	2010	2011	2012	2013	2014
# Applicants	6828	1726	183	415	407	371	347
Income ($)	42,219	18,477	17,138	17,971	18,718	18,191	19,835
2 Bedrooms	76.5%	69.8%	69.9%	68.9%	69.8%	68.2%	72.6%
3 Bedrooms	23.4%	29.8%	28.4%	30.8%	30.0%	31.8%	27.1%
Lives in Cambridge	49.3%	57.4%	61.7%	55.2%	62.4%	52.6%	57.1%
Works in Cambridge	55.2%	39.7%	28.4%	36.6%	39.8%	44.7%	44.1%
Age Youngest Member	10.5	8.5	8.2	8.2	8.2	8.7	9.0
Age Oldest Member	40.0	36.7	34.7	35.7	36.6	37.7	37.7
# Children	1.25	1.27	1.25	1.39	1.27	1.24	1.16
Child Under 10	60.8%	60.8%	56.8%	56.6%	62.9%	62.0%	64.8%
Household Head Head White	55.2%	36.2%	37.2%	32.3%	38.8%	38.8%	34.3%
Household Head Head Black	19.6%	50.3%	55.7%	54.7%	47.7%	46.6%	49.3%
Household Head Head Hispanic	17.9%	19.2%	17.5%	20.2%	17.2%	20.8%	19.9%

Notes: The applicant sample consists of Family 2-3 bedroom priority applicants who made their initial development choices between 2010 and 2014. Application date is defined as the first date an applicant appears on a waitlist in the status log. Family Public Housing waitlists were closed during the second and third quarters of 2010. The eligible population is estimated using the 2010-2014 American Community Survey (ACS). Households already living in Cambridge public housing, as well as households that applied before 2010 and were still on the waitlist during the sample period, are not counted as eligible.

Table 3: Characteristics of Family Public Housing Developments

List Name	Mean Waiting Time (Years)	# Housed Applicants	# Units	Neighborhood	Tenant Income ($)	Applicant Income ($)	African American Household Head
Roosevelt Mid-Rise	1.58	18	77	East	18,370	13,930	41%
Woodrow Wilson	1.98	2	68	Central	21,181	15,662	75%
Jefferson Park	2.16	62	284	North	27,982	16,025	62%
Newtowne Court	2.33	95	268	Central	23,368	16,619	62%
Washington Elms	2.92	26	175	Central	31,795	16,237	61%
Putnam Gardens	2.98	36	122	Central	22,460	16,896	60%
Corcoran Park	3.05	45	153	North	26,968	17,923	65%
Scattered	3.52	11	88	N/A	25,480	17,064	63%
Roosevelt Low-Rise	3.55	21	124	East	28,929	18,040	63%
Lincoln Way	3.72	2	70	North	32,528	17,960	62%
Jackson Gardens	3.75	9	45	Central	22,352	17,322	47%

Notes: Characteristics of CHA Family Public Housing developments available between 2010 and 2014. Mean Waiting Time is the mean waiting time for applicants who were housed during the sample period. Tenant characteristics reflect active tenant certifications on January 1st, 2014. Applicant characteristics reflect all applicants who selected the list as an initial choice. The "Scattered" list aggregates three lists: Mid Cambridge, East Cambridge, and River Howard Homes.

Table 4: Initial Development Choices

Applicant Group	Number of Applicants	Selectivity: 2 Initial Choices	3 Initial Choices	Mean Waiting Time (Years)	Number of Units	Location: # Central Cambridge	# East Cambridge	# North Cambridge
All	1726	12.1%	84.1%	2.89	145	1.50	0.51	0.79
Panel A: Household Income								
$0 - 8,000	466	11.2%	85.0%	2.86	148	1.50	0.52	0.79
$8,000 - 16,000	411	10.7%	85.6%	2.87	145	1.51	0.54	0.77
$16,000 - 32,000	555	10.8%	85.2%	2.89	145	1.50	0.50	0.82
Over $32,000	294	17.7%	78.2%	2.98	142	1.48	0.49	0.77
Panel B: Neighborhood of Current Residence								
Central Cambridge	521	9.8%	85.8%	2.89	141	1.68	0.50	0.63
East Cambridge	131	12.2%	84.0%	2.94	136	1.46	0.87	0.47
North Cambridge	338	19.2%	76.9%	2.93	147	1.26	0.37	1.11
Outside Cambridge	736	10.3%	86.1%	2.87	150	1.49	0.52	0.82

Notes: Characteristics of initial choices, by applicant characteristics. Initial choice characteristics are first averaged across each applicant's chosen developments, and then averaged across applicants. Sample is Family 2-3 bedroom priority applicants who made their initial choices between 2010 and 2014. Neighborhood is based on the zip code of the applicant's contact address. East contains zip codes 02141 and 02142; Central contains 02139; North contains 02138 and 02140; and Outside Cambridge contains all other zip codes.

Table 5: Final Development Choice

	No Controls		Development Controls		Choice Set Controls	
	(1)	(2)	(3)	(4)	(5)	(6)
Position on Waiting List	-0.0175		-0.0191		-0.0259	
	(0.0031)		(0.0036)		(0.0063)	
Expected Waiting Time (Years)		-0.0639		-4.051		-4.992
		(0.279)		(0.755)		(1.319)
Development FE's			X	X		
Development - Choice Set FE's					X	X
Implied Own-Price Elasticity	-0.657	-0.029	-0.747	-3.511	-1.125	-4.087
	(0.145)	(0.128)	(0.175)	(0.669)	(7.677)	(2.121)
Observations	573	573	573	573	343	343

Notes: Estimates from a conditional logistic regression of final development choice on waiting time information from the applicant's final choice letter. Sample is applicants who made a final development choice between 2010 and 2014. List position is calculated for each applicant/list on the date the Cambridge Housing Authority sent the final choice letter. Continued waiting time is estimated from realized waiting times after applicants made their final choices. Columns (1) and (2) have no controls. Columns (3) and (4) include fixed effects for each development. Columns (5) and (6) include as fixed effects a full set of interactions between the development and the applicant's choice set.

Table 6: Parameter Estimates

	Baseline Specification (1)		Richer Observed Heterogeneity (2)		Unobserved Taste for Size and Location (3)	
Annual Discount Rate	0.973	(0.015)	0.975	(0.013)	0.972	(0.015)
S.D. Development Fixed Effects	0.263		0.167		0.131	
Panel A: Value of Assistance						
Head Is Black	0.689	(0.047)	0.379	(0.046)	0.395	(0.044)
Head Is Hispanic	-0.011	(0.1)	-0.163	(0.068)	-0.194	(0.112)
Lives In Cambridge	0.487	(0.048)	0.210	(0.032)	0.201	(0.04)
Youngest Member < 10 Years			0.293	(0.038)	0.293	(0.04)
3 Bedroom Household			0.079	(0.054)	0.080	(0.031)
Household Income < $20,000			0.379	(0.05)	0.378	(0.095)
Log Of Observed Income	0.011	(0.047)	0.234	(0.047)	0.231	(0.053)
Log Of Observed And Unobserved Income	-1.000	--	-1.000	--	-1.000	--
Scale of R.E. Unknown Income ($10,000)	1.931	(0.09)	1.716	(0.091)	1.719	(0.127)
Panel B: Match Values						
Applicant and Development Same Neighborhood	0.020	(0.084)	-0.216	(0.047)	-0.249	(0.068)
Applicant Head Is Hispanic * Development Size			0.128	(0.04)	0.126	(0.037)
Youngest Member < 10 Years * Development Size			0.039	(0.015)	0.031	(0.018)
Household Income < $20,000 * Development Size			-0.025	(0.021)	-0.022	(0.02)
S.D. Unobserved Taste For Development Size					0.031	(0.013)
S.D. Unobserved Taste for North Cambridge					0.031	(0.012)
S.D. Unobserved Taste for East Cambridge					0.005	(0.004)
S.D. Idiosyncratic Shock	0.152	(0.01)	0.124	(0.007)	0.124	(0.01)

Table 7: Equivalent Variation to Moving from Lower-Ranked to 1st Choice Development

		All Applicants		African American Household Head		Household Income Below $15,000	
		Median	Mean	Median	Mean	Median	Mean
1st Choice instead of 2nd	% Income	7.2	9.5	7.2	9.6	7.3	9.6
	$/year	563	1,195	610	1,374	438	897
1st Choice instead of 3rd	% Income	14.2	16.3	14.3	16.3	14.3	16.3
	$/year	1,158	2,034	1,297	2,361	842	1,518
1st Choice instead of Last	% Income	152.9	158.8	155.4	160.2	155.4	161.4
	$/year	14,171	19,482	17,300	23,006	8,509	15,114

Notes: Equivalent variation of re-assigning applicants from a less preferred development to their first choice, averaged across a simulated sample of eligible households that would apply for Cambridge public housing. The simulation uses estimates from Specification (3). % Income is the difference in equivalent variation as a percentage of the household's observed income.

Table 8: Willingness to Accept Mismatched Offers

Number of Acceptable Developments	All Applicant Households			African American Head			Income below $15,000		
	%	Outside Option ($)	Observed Income ($)	%	Outside Option ($)	Observed Income ($)	%	Outside Option ($)	Observed Income ($)
1	6.7	25,665	30,251	5.3	33,409	31,616	3.3	24,872	8,454
2	4.8	23,522	29,434	3.8	29,877	29,343	1.9	22,840	8,546
3	3.6	23,150	28,194	3.7	28,747	30,006	1.8	22,926	8,607
4	3.2	20,835	28,325	2.4	27,050	23,725	1.9	20,799	7,610
5	3.0	21,328	24,191	2.9	27,608	24,815	2.2	20,973	8,408
6	3.0	19,207	26,594	2.3	24,878	25,661	1.9	19,088	9,314
7	3.1	18,826	22,863	2.6	23,936	21,157	2.4	18,279	8,225
8	3.2	18,547	23,869	2.6	24,268	23,256	2.4	17,775	8,929
9	3.4	17,034	21,521	3.8	20,044	22,815	2.9	16,330	8,185
10	4.4	15,264	21,419	4.7	18,508	18,909	3.5	14,536	7,919
11	5.2	14,252	20,955	5.4	17,945	17,451	4.0	14,215	8,594
12	6.3	12,599	19,302	7.7	15,379	17,366	5.8	12,296	7,973
13	50.1	5,122	14,489	52.7	5,931	12,342	66.0	4,828	7,167

Notes: Distribution of number of acceptable developments, averaged across a simulated sample of eligible households that would apply for Cambridge public housing. The simulation uses estimates from Specification (3). Outside Option Value is the household's budget constraint outside of public housing, including unobserved income.

Table 9: Effects of Alternative Development Choice Systems under Equal Priority

	Choose One	Choose Any Subset	Choose All or One	Choose Neighborhood	Choose All or Neighborhood	No Choice
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Welfare Gain and Cost of Allocation						
Equivalent Variation ($)	15,243	15,239	15,078	12,024	12,029	10,406
Cost per Unit ($)	18,524	18,520	18,525	18,909	18,919	18,929
Equivalent Variation per $ Cost to Gvt.	0.82	0.82	0.81	0.64	0.64	0.55
Panel B: Targeting						
Observed Income ($)	18,252	18,268	18,250	16,971	16,935	16,903
Observed and Unobserved Income ($)	9,947	9,930	9,949	8,541	8,528	8,378
% Extremely High-Need	44.6%	44.8%	44.6%	50.0%	50.0%	46.9%
Panel C: Match Quality						
% Assigned Top Choice	52.3%	51.2%	52.3%	15.9%	15.8%	7.7%
% Assigned Top 3	76.6%	76.5%	76.5%	36.1%	36.0%	22.1%
Panel D: Characteristics of Housed Applicants						
Waiting Time (days)	1433	1438	1433	830	830	786
% Black	42.2%	42.1%	42.2%	44.4%	44.5%	42.6%
% Hispanic	21.4%	21.4%	21.4%	16.0%	16.1%	17.9%
From Cambridge	65.0%	65.1%	65.0%	62.9%	62.9%	64.6%

Notes: Statistics averaged across assigned apartments in each counterfactual simulation. Dollar amounts are annual. Cost per Unit is calculated based on market rate rental prices in Cambridge, MA during the sample period. Equivalent Variation is the equivalent cash transfer outside of public housing that would generate the same welfare change for a housed applicant as their assignment. Extremely High-Need applicants are at the minimum consumption level of $10/day outside of public housing.

Table 10: Alternative Choice and Priority Systems

	Common Choice and Priority Systems						Full Information	
	Low-Income Priority		High-Income Priority		Equal Priority		Equivalent Variation Maximizing	Targeting Maximizing
	Choose One	No Choice	Choose One	No Choice	Choose One	No Choice		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Welfare Gain and Cost of Allocation								
Equivalent Variation ($)	15,444	11,274	12,944	9,651	15,243	10,406	18,087	13,319
Cost per Unit ($)	19,948	21,240	17,302	16,914	18,524	18,929	19,665	19,875
Equivalent Variation per $ Cost to Gvt.	0.77	0.53	0.75	0.57	0.82	0.55	0.92	0.67
Panel B: Targeting								
Observed Income ($)	13,505	9,200	22,326	23,619	18,252	16,903	14,451	13,749
Observed and Unobserved Income ($)	8,202	7,183	10,711	9,473	9,947	8,378	6,381	4,041
% Extremely High-Need	51.9%	56.6%	40.0%	38.2%	44.6%	46.9%	61.7%	80.0%
Panel C: Match Quality								
% Assigned Top Choice	41.6%	7.0%	45.1%	8.7%	52.3%	7.7%	23.5%	6.6%
% Assigned Top 3	63.8%	21.4%	62.8%	23.4%	76.6%	22.1%	43.9%	19.8%
Panel D: Characteristics of Housed Applicants								
Waiting Time (days)	850	289	1075	443	1433	786	101	74
% Black	44.7%	48.0%	40.0%	38.2%	42.2%	42.6%	58.6%	37.9%
% Hispanic	21.6%	19.5%	18.8%	16.4%	21.4%	17.9%	12.5%	21.0%
From Cambridge	65.1%	67.8%	62.4%	62.1%	65.0%	64.6%	69.9%	64.0%

Notes: statistics averaged across assigned apartments in each counterfactual simulation. Dollar amounts are annual. Cost per Unit is calculated based on market rate rental prices in Cambridge, MA during the sample period. Equivalent Variation is the equivalent cash transfer outside of public housing that would generate the same welfare change for a housed applicant as their assignment. Extremely High-Need applicants are at the minimum consumption level of $10/day outside of public housing. Low-Income Priority first offers vacant apartments to applicants with incomes below $15,000; High-Income Priority does the same for applicants with incomes above $15,000.

Figure 1: Welfare Under Alternative Development Choice Systems

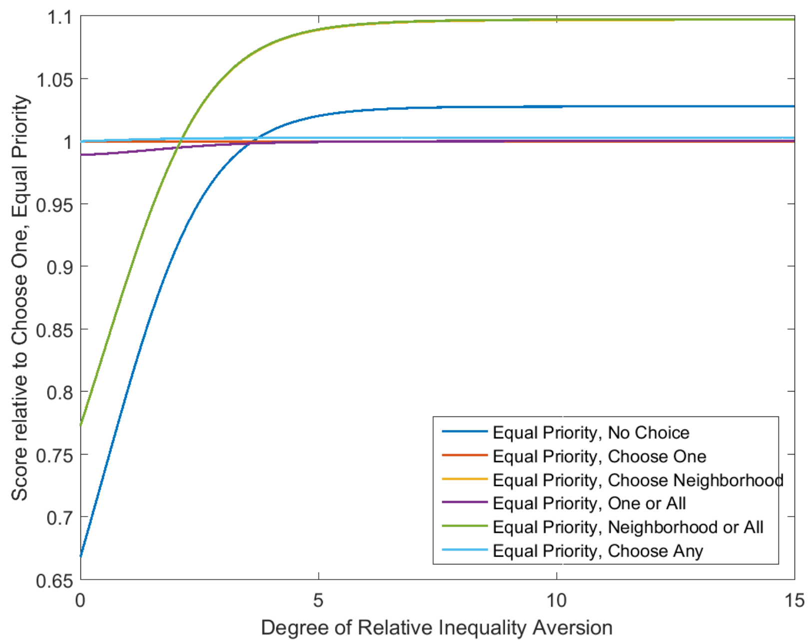
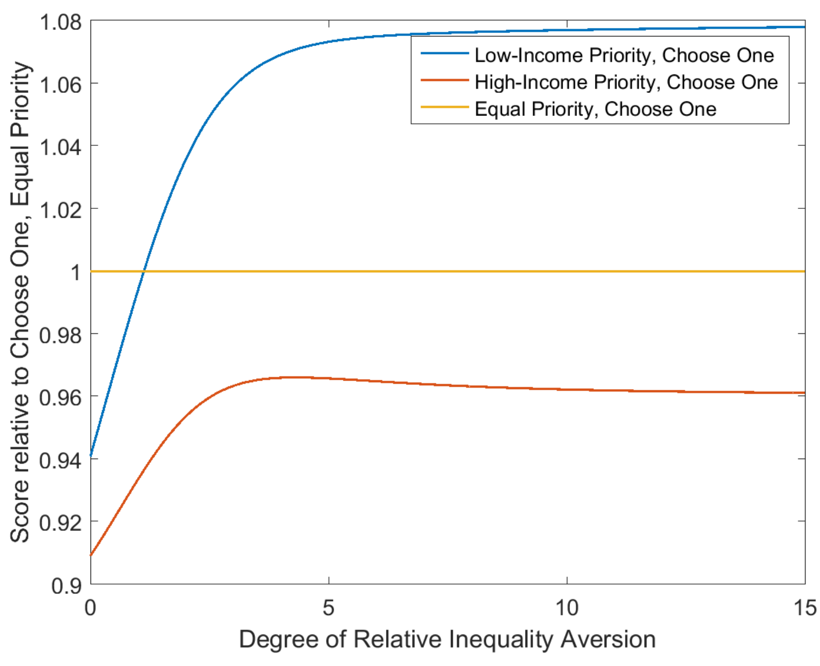


Figure 2: Welfare Under Alternative Priority Systems



Notes: Figures 1 and 2 compare cost-adjusted welfare gains produced by different choice and priority systems used in practice, defined in Section 7.1. Figure 1 compares alternative choice systems under Equal Priority. Figure 2 compares alternative priority systems under the Choose One development choice system. Each point on the x-axis corresponds to a degree of relative inequality aversion. At each point, cost-adjusted welfare gains from each mechanism are normalized by the value for Equal Priority, Choose One.

Figure 3: Preferred Mechanisms by Degree of Inequality Aversion

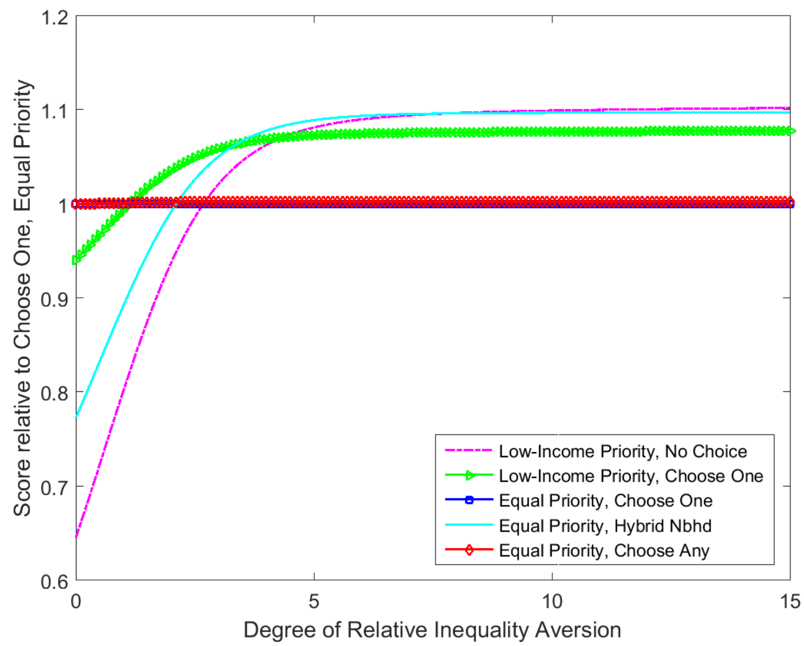
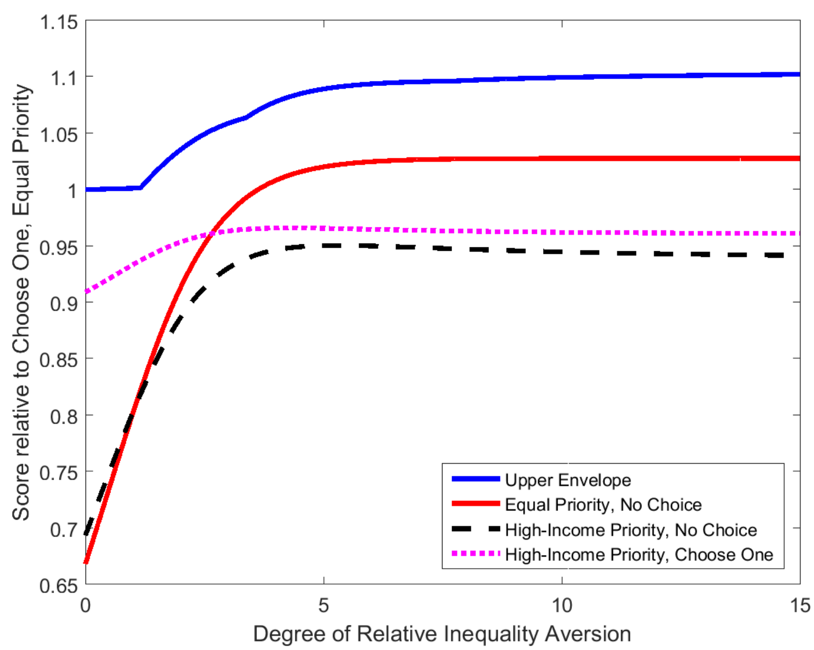


Figure 4: Mechanisms Dominated in Constant Relative Inequality Aversion Class



Notes: Figures 3 and 4 plot cost-adjusted welfare gains from choice and priority systems as a function of the degree of relative inequality aversion. Figure 3 plots the welfare effects of mechanisms which perform best for some degree of inequality aversion. Figure 4 plots the upper envelope from the first subfigure along with three mechanisms that never perform best for any degree of inequality aversion. Each point on the x-axis corresponds to a degree of relative inequality aversion. At each point, cost-adjusted welfare gains from each mechanism are normalized by the value for Equal Priority, Choose One.

For Online Publication: Appendix to "Targeting In-Kind Transfers Through Market Design: A Revealed Preference Analysis of Public Housing Allocation"

Daniel Waldinger

April 3, 2019

[*]I am very grateful to Nikhil Agarwal, Parag Pathak, Michael Whinston, and Amy Finkelstein for invaluable guidance and support. I also thank Nick Arnosti, Glenn Ellison, Ingrid Gould-Ellen, John Lazarev, Jacob Leshno, Peng Shi, Neil Thakral, and Heidi Williams; seminar participants at the MIT Industrial Organization and Public Finance workshops; my colleagues at NYU; and participants in many other seminars and conferences. The Cambridge Housing Authority generously provided the applicant and tenant data used in this paper, with special thanks to Tara Aubuchon, Tito Evora, Michael Johnston, Jay Leslie, Hannah Lodi, and John Ziniewicz. All analysis and views expressed in this paper are my own and do not represent the views of the Cambridge Housing Authority. I acknowledge support from a National Science Foundation Graduate Research Fellowship. All mistakes are my own.

[†]NYU Furman Center. Email: danielwaldinger@nyu.edu.

A Datasets

A.1 CHA Dataset and Sample Selection

The Cambridge Housing Authority maintains a database of applicants and tenants to manage its programs and comply with HUD regulations. The dataset used in this paper is based on an extract made on February 26th, 2016. It contains anonymized records of all applicants for Cambridge public housing who were active on a waiting list between October 1st, 2009 and February 26th, 2016. This includes all households who submitted an application after October 2009, and a selected sample of households who applied before late 2009 and were still on the waiting list.

For each applicant, I observe household characteristics, development choices, and the timing and outcome of all events during the application process. Household characteristics include family size; the age, gender, and race/ethnicity of each household member; zip code of current residence; and self-reported household income. The data also record whether an applicant had priority. Development choices and waiting list events come from a time-stamped status log that records the status of each application over time. This includes the applicant's initial application date; the date it joined each waiting list; the date it was sent a final choice letter, and if it responded, its final choice; and the date the applicant was offered an apartment. I also observe the date and reason if a household was removed from the waiting list.

From the application data, I construct several objects that allow me to interpret development choices. I infer the set of developments for which each applicant was eligible based on household structure and application date.[1] I observe waiting times for applicants who were offered apartments, both from initial application and from the date the applicant made its final choice. I also infer the information each applicant received in their final choice letter by computing the applicant's list position on the date CHA sent the letter.

For analysis, I restrict my sample to priority applicants for 2 and 3 bedroom apartments in the Family Public Housing program who submitted an application between January 1st, 2010 and December 31st, 2014. Non-priority applicants had virtually no chance of being housed, so it is unclear how to interpret their development choices. Family Public Housing applicants are a more homogeneous group than Elderly/Disabled households, and families with children are of substantial policy interest. I restrict to 2 and 3 bedroom apartments for sample size; the vast majority of Family Public Housing applicants apply for these units, and data on choices, waiting times, and list positions from each development are sparse for other bedroom sizes. Analyzing new applications between 2010 and 2014 avoids selection issues with pre-2010 applicants since some pre-2010 applicants were no longer on the waiting list at the beginning of the sample period. These restrictions produce a sample of 1,752 applicants. 26 of these applicants selected more than three developments; omitting them leaves 1,726 applicants for structural

[1] To reduce waiting time uncertainty, CHA merged four small waiting lists with larger lists in 2013. As a result, an applicant's initial choice set depended on its application date.

estimation.

A.2 American Community Survey

The American Community Survey (ACS) publishes anonymized, household-level micro-data covering 1 percent of the U.S. population each year. The years 2010-2014 form a 5 percent sample of U.S. households. The survey collects detailed information on each household's structure, geography, and economic and demographic characteristics. Data can be downloaded at `https://usa.ipums.org/usa-action/variables/group`.

The ACS contains key household-level information that determines whether a household could have appeared in my applicant sample, which contains applicants with priority for 2 and 3 bedroom apartments in Cambridge Family Public Housing. I begin with the universe of ACS households living in the state of Massachusetts. I then determine whether each household lived or worked in Cambridge.[2] Cambridge has its own city code since its population is greater than 100,000. The $CITY$ field identifies whether each household lives in Cambridge, and place of work for each working household member comes from the $PWPUMA00$ field. To determine a household's bedroom size, I apply the rule used by the CHA based on the age and gender of each household member and their relation to the household head. I also identify whether households would have been eligible for the Elderly/Disabled or the Family Public Housing program based on the age of the oldest household member. For households composed of three or more generations, I created separate households for the elderly members and the younger members.[3] For income eligibility, I divide the household's total income by the Area Median Income for their household size and survey year. Other characteristics of eligible ACS households, such as the race, ethnicity, and gender of the household head, are determined using ACS demographic variables.

[2] There are tens of thousands of households with veteran status in Massachusetts, so veteran status is not counted to determine which households would have had priority for Family Public Housing in Cambridge. Only a small number of applicants have veteran status, and most already live in Cambridge.

[3] According to the CHA, it is common for Family Public Housing applicants to apply with a two-generation subset of their current multi-generational household.

B Robustness Checks for Final Choice Analysis

For the evidence of responsiveness to waiting time information in section 3.3 to be valid, position information provided to applicants when they make their final choice should be uncorrelated with their preferences, conditional on first-stage decisions. While difficult to test directly, the data rule out two possibilities that would suggest this condition is violated. First, conditional on an applicant's initial choice, the position information they receive at final choice is uncorrelated with their observable characteristics. If there is selection into the final choice stage based on preferences, it is only a function of unobservables. Second, in contrast to the final choice analysis, initial choices are not predicted by list lengths on the specific date a household applied. This suggests that applicants were not aware of short- or medium-term fluctuations in list lengths before they received their final choice letters.

B.1 Testing for Selection in Final Choice Analysis

While one cannot directly test random assignment of final choice list positions, one can check whether these positions are correlated with observed characteristics. The idea is analogous to a test for balance between treatment and control group characteristics in a traditional randomized controlled trial. Here, the analysis is complicated by two factors. First, because each applicant selected their final choice set in the first stage, the test must condition on initial choices. Second, the "treatment" is multi-dimensional because each applicant learns up to three list positions at the final choice stage.

I therefore test whether each *pair* of list positions in the applicant's final choice letter predicts the applicant's characteristics. Let C_i denote applicant i's initial choice, p denote a pair of developments $j, k \in C_i$, and Z_i denote an applicant characteristic. I run an ordinary least squares regression with one observation for each applicant and pair of developments in a final choice letter:

$$Z_i = \alpha_{p,C_i} + \beta_{p,C_i} \frac{x_{ij}}{x_{ik}} + \epsilon_{ip} \tag{22}$$

The predictors include choice set, pair interaction dummies and the ratio of list lengths between j and k quoted to applicant i. This ratio can have a different relationship with the dependent variable for each choice set, pair interaction. One can interpret the ratio as a relative price; a higher value means that the applicant faces a longer continued wait for development j relative to development k. To test whether certain types of applicants systematically receive different position information at final choice, captured by these relative prices, I perform a joint F-test of the hypothesis $\beta_{p,C_i} = 0 \; \forall p, C_i$.

Table 11 shows that final choice list positions are not significantly predictive of most applicant characteristics. Panel A constructs the relative price using list position, while Panel B uses expected continued waiting time. The first two characteristics – application date and final choice letter date – are strongly predicted by final choice positions. These relationships are to be expected, and they demonstrate that the regressions have sufficient power to reject the null hypothesis. The correlation between application date and position information is consistent with the fact that some lists are becoming longer relative to others over the sample period. The correlation between final choice date and position information would occur for purely mechanical reasons, even without trends in relative list lengths over time. If an applicant receives a final choice letter early compared to others who made the same initial choice, one of the lists must be unusually short.

Among the other characteristics, only number of children and number of household members, which are highly (and mechanically) correlated, have F-test p-values below .05. Importantly, characteristics that are important predictors of applicant behavior in the structural model, including household race/ethnicity and annual income, are not correlated with position information conditional on initial choices. If applicants were selecting into the final choice stage based on their development preferences, one would expect selection to be correlated with these characteristics. The absence of a correlation supports interpreting the final choice regressions in section 3.3 as revealing a causal response of final choices to waiting time information.

B.2 Testing for Responsiveness of Initial Choices to List Position

To test for responsiveness of initial choices to list position, I construct a dataset similar to the one used for the final choice analysis in section 3.3. Specifically, for each applicant, I determine the position they would be on the waiting list for each development if they included that development in their initial choice. I then test whether each development is more likely to be selected on dates when that development's waiting list is short relative to those of other developments. Conducting this analysis separately for each development deals with the fact that an applicant may select multiple developments in their initial choice.

Define $y_{ij} = 1\{j \in C_i\}$ to be an indicator for whether applicant i selected development j as part of their initial choice, and let x_{ij} be the position number the applicant would have had on list j if they selected it (regardless of whether they actually did). Each applicant has characteristics Z_i. I estimate the following regression equation for each development j:

$$y_{ij} = \alpha_{ij} + \delta Z_i + \sum_{k \neq j} \beta_{jk} \frac{x_{ik}}{x_{ij}} + \epsilon_{ij} \tag{23}$$

Equation 23 allows the probability development j is selected to depend on the ratio between the length of list j and list k for every other development k. This specification captures the idea that *relative* list lengths should matter for applicants' decisions, and also allows applicant characteristics to predict their choices.

Table 12 presents F-statistics and p-values from a test of the joint hypothesis $\beta_{jk} = 0 \; \forall k$ for each development j. In Panel A, which controls for list lengths but not applicant characteristics, three developments have p-values below 0.05. Only one development has a p-value below 0.05 once applicant characteristics are included in Panel B. There is therefore little evidence of responsiveness to list position at the initial choice stage. While these patterns contrast starkly with the clear response to position information at final choice, they are consistent with institutional facts about the Cambridge Mechanism. The CHA did not make list length information readily available to new applicants, and although an applicant could call the CHA and ask for its position number on each list after it applied, few did so.

In addition to validating the final choice analysis by ruling out a particular source of selection into the final choice stage, table 12 also motivates the information structure in the development choice model. Applicants do not behave as though they know the length of each list when they apply; instead, their initial choices are consistent with a common prior based on a steady state distribution of waiting times, while their final choices show updating based on the specific position information in their final choice letters.

C Estimation Details

C.1 Waiting Time Beliefs

This section provides details of the simulation-based procedure to estimate applicant beliefs using knowledge of the Cambridge Mechanism and waiting list data. Since applicants choose developments in two stages, select multiple developments in the first stage, and make choices based on new information in the second stage, the waiting lists for different developments move interdependently. A sophisticated applicant will account for the fact that the combination of developments selected in the first stage will jointly affect the conditions under which they make their final development choice in the second stage. They will also update their beliefs about continued waiting times given their positions on all three lists at the final choice stage. This poses a challenge for estimation since data on realized waiting times given initial choices and final choice states are sparse. A parsimonious model of dependence across lists may not be realistic.

I assume that beliefs are consistent with the steady-state distributions that the Cambridge Mechanism would generate given applicant arrival and departure rates, initial and final choice frequencies, and empirical vacancy rates. These empirical quantities can be estimated directly from application data. Combining these estimates with knowledge of the Cambridge Mechanism, I simulate steady state outcomes which quantify interdependence across lists and the option value of the timing and information of the final choice stage.

C.1.1 Cambridge Mechanism

Between 2010 and 2014, Cambridge ran its public housing waiting lists according to the following algorithm. Calendar time is indexed $t = 1, ..., T$. Waiting lists are indexed by $j = 1, ..., J$, where a list corresponds to a specific bedroom size apartment (2 or 3 bedrooms) in a specific development. Applicants are indexed $i = 1, ..., N$, vacancies by $\nu = 1, ..., V$. Applicant i has an arrival date t_i and a latent departure date r_i, and makes initial choice C_i. Vacancy ν occurs on date t_ν on list j_ν. For each list j, there is a sequence of trigger and batch size policies $\{(L_{j,k}, K_{j,k})\}_{k=1}^{K}$ for sending final choice letters. If fewer than $L_{j,k}$ applicants on list j have made a final choice, Cambridge sends final choice letters to the next $K_{j,k}$ applicants on list j who have not yet made a final choice. The pair $(L_{j,k+1}, K_{j,k+1})$ become the next trigger and batch policy for list j. x_{ij} is applicant i's list j position in its final choice letter, computed as the total number of applicants on list j with an earlier application date on the date the letter is sent. Finally, the coefficients for the final choice model are $(\beta, \{\xi_j\}_{j=1}^{J})$.

The Cambridge mechanism proceeds as follows. The simulation begins at $t = 0$ with empty lists, no vacant units, and an initial trigger and batch policy $(L_{j,1}, K_{j,1})$ for each list. The following occurs in each period t:

(i) Each applicant i with arrival date $t_i = t$ is added to the lists in its initial choice set ($j \in C_i$).

(ii) Each vacancy ν with $t_\nu = t$ is offered to the first applicant on list j_ν who has made a final choice. Applicant i is housed in j_ν and removed from the waiting list. If no applicants are available, the vacancy is pushed to next period (t_ν is moved to $t_\nu + 1$).

(iii) For each list j, if the number of applicants who are on list j and have made their final choice is less than the current trigger $L_{j,k}$, the following steps occur:

   (a) Cambridge sends final choice letters to the first $K_{j,k}$ applicants on list j who have not made their final choice.

   (b) Applicant i responds to the final choice letter if $r_i \geq t$

   (c) If i responds, it chooses list j with probability

$$\frac{\exp(\beta x_{ij} + \xi_j)}{\sum_{m \in C_i} \exp(\beta x_{im} + \xi_m)}$$

   (d) If i does not respond, it is removed from all lists $m \in C_i$

   (e) The next trigger and batch policy, $(L_{j,k+1}, K_{j,k+1})$, is drawn for next period

   Otherwise, $(L_{j,k}, K_{j,k})$ is held for the next period.

(iv) Each applicant with $t_i = t$ who has already made its final choice is removed from the list.

C.1.2 Inputs to Simulation

Simulation of the Cambridge Mechanism requires a sequence of applicant arrival dates t_i and the initial choice C_i and departure date r_i of each arrival; a sequence of apartment vacancies with dates t_ν on list j_ν; and a sequence of batch and trigger policies $\{L_{j,k}, K_{j,k}\}_{k=1}^{K}$ for each list j. I assume that all sequences are drawn independently and make the following parametric assumptions:

- Applicants arrive at a poisson rate α
- Each applicant departs immediately with a non-zero probability a_1 and at exponential rate a_2 after.
- Applicant choices are drawn uniformly from the empirical distribution in the Cambridge dataset
- Vacancies on each list occur at poisson rate $v_j = 0.1 * S_j$, where S_j is the number of units corresponding to list j. The sequences occur independently across developments and bedroom sizes.
- The sequence of trigger and batch policies is drawn with uniform probability from its empirical distribution in the Cambridge dataset.
- Final choice probabilities are determined by Specification (3) in Table 4, in which the latent utility of each option depends on list position and a development fixed effect.

Given these primitives, I draw inputs for a 500 year simulation and run the Cambridge mechanism. Waiting times converged after about 10 years. I used the last 490 years of the simulation to construct beliefs.

C.1.3 Constructing Belief Objects

The simulation produces the state of all Cambridge waiting lists every day for 490 years. To estimate the relevant distributions governing beliefs, I consider what would have happened to an additional applicant arriving on each simulation date, for each sequence of choices the applicant could have made.

To estimate $\{G_C(S_C, P_C)\}_{C \in \mathcal{C}}$, the distribution of final choice states after making each initial choice C, I sample 1000 dates $t_1, ..., t_{1000}$ from the simulation. For every C, I compute the date s_C and position vector p_C that an applicant who applied on date t_s would have received, for $s = 1, ..., 1000$. These states – $\{(s_C^s, p_C^s)\}_{s=1,...,1000}$ – form an empirical measure $\hat{G}_C$.

Constructing beliefs $\{F_{j,C}(. \mid p_C)\}_{j,C,p_C}$ for continued waiting time at final choice is more complicated. There are over 1800 possible (j, C) initial and final choice combinations, and for each combination, each position vector p_C induces a different continued waiting time distribution. Even using the simulation results, there is a limit to how flexibly these distributions can (and should) be estimated. My approach is to specify a hierarchical parametric model for the continued waiting time distribution. I assume that continued waiting time follows a beta distribution

$$T_j \mid j, C, p_C \sim Beta(\alpha_{j,C}(p_C), \beta_{j,C}(p_C))$$

whose parameters depend flexibly on choices j and C and parametrically on positions p_C. For a (j, C) pair with $|C| = 3$, the position vector p_C enters the beta distribution parameters as

$$\alpha_{j,C}(p_C) = \exp\{\pi_1 p_1 + \pi_2 \log(p_1) + \pi_3 \log(p_2) + \pi_4 \log(p_3)\}$$

$$\beta_{j,C}(p_C) = \exp\{\pi_5 p_1 + \pi_6 \log(p_1) + \pi_7 \log(p_2) + \pi_8 \log(p_3)\}$$

where the π parameters are (j, C)-specific. p_1 is the position on list j, and p_2 and p_3 are the other positions. I found that this parametric specification did a good job fitting the distribution of realized waiting times from the simulation. The range of each beta distribution is $[0, \lceil \max T_{j,C} \rceil]$.

The hierarchical parameters of each beta distribution are estimated as follows: for computational speed, I take a 5% sample of application dates from the simulation denoted $\{t_d\}_{d=1,...,D}$. For each initial choice C, I calculate the position vector an applicant would have received in their final choice letter, as well as the continued waiting time for each list. From this dataset of position vectors and continued waiting times $\{p_{C,d}, t_{C,d}\}_{d=1,...,D}$, π and the upper bound of the support of the beta distribution for each $j \in C$ are estimated by maximum likelihood.

C.2 Development Preferences

C.2.1 Moments

To estimate the parameter vector $\theta = \{\rho, \delta, \beta, g(.), \sigma_\eta\}$, I match the following sets of moments:

- Application rates by income and demographics: I currently use the following characteristics Z_i, all of which are indicator variables: a dummy equal to 1 for all households; annual household income in the ranges of $[X, X + 10,000]$ for X in \$10,000 intervals from \$0 to \$50,000; the household head is black or hispanic; the household currently lives in Cambridge; the youngest household member is less than 10 years old; the household requires three bedrooms; and household income is below \$20,000. I also match the rate at which all households and households earning \$0-\$20,000 and \$20,000-\$40,000 select three developments in their initial choice.
- Development shares: there is one moment for the initial choice share of each of the thirteen developments.
- Covariances between applicant characteristics and characteristics of their initial development choices. I match the rates at which Cambridge residents select developments in their current neighborhood of residence, and the covariance between chosen development size and whether the household head is hispanic, the household requires three bedrooms, and the household's youngest member is less than 10 years old.
- Means and Variances of chosen development characteristics within and between applicants. Each of these moments is constructed for development size (# units) and whether the development is in North, East, or Central Cambridge. For households that do not apply, all moments are zero.
- Means and variances of chosen waiting times within and between applicants, by income and demographics. The first and second waiting time moments are interacted with household income bins for \$0-\$20,000, \$20,000-40,000, and \$40,000+.
- The final choice moments used are:
  - The fraction of eligible households who made a final choice:

$$m_i^{(q)} = 1\{f_i \neq \emptyset\}$$

  - The mean expected *continued* waiting time of final choices, given an applicant's position information:

$$m_i^{(q)} = 1\{f_i \neq \emptyset\} t_{f_i}$$

  - The *relative price index*, as an expected continued waiting time ratio, of the final choice compared to other developments in each applicant's choice set. If $C = \{j, k, m\}$, and the expected continued waiting times for the developments are $\{t_j, t_k, t_m\}$, then the relative price index for development j is defined

$$R_{j,C} = \frac{1}{2}\left[\frac{t_j}{t_k}/\bar{r}_{jk,C} + \frac{t_j}{t_m}/\bar{r}_{jm,C}\right]$$

where $\bar{r}_{jk,C}$ is the mean continued waiting time ratio between developments j and k for applicants who made a final choice from choice set C. The resulting moments are

$$m_i^{(q)} = 1\{f_i \neq \emptyset\}R_{f_i,C_i}, \qquad 1\{f_i \neq \emptyset\}1\{R_{f_i,C_i} > 1\};$$

The relative price index captures whether an applicant faced a high or a low "price" for its final choice f_i, compared to other applicants who made their final choice *from the same choice set* C_i. This isolates the natural experiment created by the Cambridge Mechanism, where applicants who made the same initial choices are given different waiting time information when they make their final choices.

– The average and maximum difference in expected continued waiting time between the chosen and alternative developments:

$$m_i^{(q)} = 1\{f_i \neq \emptyset\}\left(t_{f_i} - \frac{1}{2}\left[t_k + t_m\right]\right), \qquad 1\{f_i \neq \emptyset\}\left(t_{f_i} - \min\{t_k, t_m\}\right).$$

C.2.2 Simulation Procedure

I estimate the parameter vector θ based on moment conditions

$$E[(m_i - E(m_i \mid Z_i, \theta_0)) \mid Z_i] = 0\,,$$

where θ_0 is the true parameter vector, m_i contains features of household decisions, and Z_i are household characteristics. The method of simulated moments estimates $E(m_i \mid Z_i, \theta)$ by simulation. This procedure involves the following steps:

(i) For each sampled household i and simulation draws $s = 1, ..., S$,

  (a) Draw preference shocks $\{\eta_{is}, \nu_{ims}, \epsilon_{is}\}$.

  (b) For each possible initial choice C, draw the date and position information of the final choice (s_{is}^C, p_{is}^C), drawn from the distribution $G_C(S_C, P_C)$.

  (c) Draw an exogenous departure time using the attrition model. This determines whether the simulated applicant makes a final choice for a given final choice date s_{is}^C.

(ii) For each proposed value of θ and each (i, s),

  (a) Compute v_{is} according to equation 12 given z_i, θ, and the simulation draws.

  (b) Compute the optimal initial choice C_{is} according to equation 5 given v_{is}, the discount factor ρ, waiting time beliefs.

(c) If the exogenous departure date is after the final choice date, compute the applicant's final choice according to equation 1 given preferences and beliefs.

(d) Construct the conditional expectations

$$\hat{E}(m_i \mid z_i, \theta) = \frac{1}{S}\sum_{s=1}^{S} m_{is}(\theta)$$

and form moment conditions.

The one non-standard component of the simulation comes from the applicant's two-stage decision problem. Different parameter values θ will lead a simulated applicant to make different initial choices, inducing a different distribution over final choice states. I draw one final choice state for each possible initial choice and hold these draws fixed across candidate parameter values. This approach minimizes computational burden: if a simulated applicant makes same initial choice under two different parameter vectors, it makes the final choice under the same conditions.

C.2.3 Objective Function and Optimization

Because the moments used in estimation are highly correlated, the optimal weight matrix performed poorly. The estimator failed to match moments key for identifying value of assistance parameters and the discount factor, such as overall application rates and the mean waiting times of initial development choices. Instead, I used a diagonal weight matrix with elements inversely proportional to the sampling variance of the corresponding moment functions. I also placed additional weight application rates, variances of chosen development characteristics and waiting times, and final choice moments.

Minimizing the objective function was challenging because the objective function is discontinuous and not guaranteed to be convex. Fortunately, Monte-Carlo simulations suggested that a combination of global and local search consistently found a global minimum close to the true parameters. I used the following procedure: I first used MATLAB's *fmincon* function, approximating the gradient by finite differences. I found that iteratively decreasing the finite difference minimum step size, using the previous solution as a starting value, helped to ensure that the estimator searched widely while also finding a local minimum. At each local minimum, I used MATLAB's *patternsearch* algorithm to ensure that an exact local minimum was attained and to search for other local minima. I used several starting values covering a range of parameters. To limit numerical instability, the variance of each random coefficient was constrained to be less than one million.

C.2.4 Inference

The standard errors in Table 6 account for sampling error in the choices of eligible households and simulation error in constructing the simulated moments. They do not correct correct for statistical error in the minimum distance procedure used to estimate the distribution of eligible households, or for statistical error in the estimated distributions governing applicant beliefs.

The asymptotic variance of the method of simulated moments estimator is

$$(G'AG)^{-1}G'A\Omega AG(G'AG)^{-1}$$

where $G = E[\nabla_\theta g_i(\theta_0)]$, $\Omega = E[g_i(\theta_0)g_i(\theta_0)']$, and A is the symmetric positive-definite weight matrix used in estimation. For a consistent estimate of G, I evaluate the gradient of the moment functions at $\hat{\theta}$:

$$\hat{G} = \frac{1}{N}\sum_{i=1}^{N} \nabla_\theta \hat{g}_i(\hat{\theta})$$

Variance in the moment functions comes from two components: sampling error in applicant choice features m_i, and simulation error in $\hat{E}[m_i \mid z_i, \theta]$:

$$\Omega = \Omega_m + \frac{1}{S}\Omega_s$$

The empirical variance of the moment functions evaluated at $\hat{\theta}$ provides a consistent estimate of Ω_m:

$$\hat{\Omega}_m = \frac{1}{N}\sum_{i=1}^{N} \hat{g}_i(\hat{\theta})\hat{g}_i(\hat{\theta})'$$

Ω_s can be estimated consistently by

$$\hat{\Omega}_s = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{S-1}\sum_{s=1}^{S}(m_{is}(\hat{\theta}) - \hat{m}_i(\hat{\theta}))(m_{is}(\hat{\theta}) - \hat{m}_i(\hat{\theta}))'$$

where

$$\hat{m}_i(\hat{\theta}) = \frac{1}{S}\sum_{s=1}^{S} m_{is}(\hat{\theta})$$

The variance estimate is

$$(\hat{G}'A\hat{G})^{-1}\hat{G}'A\left(\hat{\Omega}_m + \frac{1}{S}\hat{\Omega}_s\right)A\hat{G}(\hat{G}'A\hat{G})^{-1}$$

C.3 Robustness to Alternative Decision Rules

C.3.1 A Naive Decision Rule

The estimation procedure outlined in sections 4 and 5 and implemented in section 6 relies on a particular model of applicants' decision rule and beliefs. The model allows applicants to be highly sophisticated – they anticipate the position information they receive in the final choice stage and understand that this information generates a portfolio choice problem at initial application. This decision rule and belief structure entail high cognitive demands for a socioeconomically disadvantaged population, in a setting where applicants have little chance to learn about the relevant distributions from experience or publicly available information.

While it is not possible to explore all possible alternative models, this section repeats the analysis in the paper using one alternative model of development choice and belief formation. In this model, applicants use a "naive" decision rule in which they do not consider the full complexity of the portfolio choice problem generated by the Cambridge Mechanism. Instead, in the initial choice stage applicants use a heuristic: they consider the value of applying for each development on its own and select the developments with the highest expected value according to this criterion. This naive choice rule rules out certain types of sophisticated behavior. For example, in a portfolio choice problem it can be optimal to select a development for its option value – even if it has a longer expected waiting time and is therefore much less likely to be eventually chosen than another slightly less desirable development, it may yield a greater increase in the value of the applicant's portfolio. It may also be optimal to omit a development in order to delay the timing of the final choice stage and obtain a later (and more precise) measure of continued waiting time.

Formally, at the initial choice stage applicants form beliefs about the marginal distribution of waiting times for each development j. Let $G_j(t)$ denote the believed probability that the waiting time for development j is less than t years. At the final choice stage, applicants form beliefs in the same way as in the sophisticated model, taking all list positions p into account when predicting the continued waiting time for development j. Let $F_{j,C}(t \mid p)$ denote the probability that continued waiting time for development j is less than t years given current list positions p. At final choice, the applicant solves the problem defined in equation 3, just as in the sophisticated model. At initial choice, the applicant solves a different problem than the one defined in equation 5:

$$\max_{C \in \{0,1,\ldots,J\}^3} \sum_{j \in C} E\left[e^{-\rho T}\right](v_{ij} - v_{i0}) \tag{24}$$

$$= \max_{C \in \{0,1,\ldots,J\}^3} \sum_{j \in C} \int \frac{1}{\rho} e^{-\rho T}(v_{ij} - v_{i0}) dG_j(T) \ . \tag{25}$$

In estimation, the distributions G_j and $F_{j,C}$ come from the same simulation that generated the belief distributions used for the main estimates. Therefore, beliefs are consistent across the two stages of choice in the sense that they are generated by a simulation that respects the structure of the Cambridge Mechanism. However, those beliefs are now inputs to a decision rule that is suboptimal because of the naive rule employed in the first stage.

I estimate the specifications in Section 6 under the naive choice rule using the same procedure as in section 5. The only difference is the decision rule and belief objects used to predict a simulated applicant's development choices in the method of simulated moments procedure. Then, I re-solve for counterfactual equilibria under alternative mechanisms using estimates obtained under specification (3) with the naive decision rule.

C.3.2 Estimates and Counterfactual Results

Alternative development choice systems produce a similar trade-off between efficiency and redistribution under the naive choice rule. In table 17, the analogue of table 9, equivalent variation per assigned unit falls from just over \$17,059 under Choose One to \$11,774 under No Choice, while the fraction of extremely high-need tenants rises from 31.9 percent to 41.2 percent. Compared to the sophisticated decision rule, the naive rule predicts slightly higher tenant values, as well as larger targeting gains from removing choice, because applicants have a higher estimated discount factor and a greater degree of estimated idiosyncratic preference heterogeneity. This leads a large fraction of applicants (over 83 percent) to select and receive their first choice development in systems that allow choice – even some very high-need applicants are sufficiently patient to wait for their preferred development. Removing choice substantially increases the probability that these applicants are housed.

Finally, relative to table 10, table 18 shows very similar comparisons across choice and priority rules assuming a naive decision rule. As under the sophisticated rule, under the naive rule cost-adjusted equivalent variation is maximized under Equal Priority, Choose One, while while the fraction of extremely high-need applicants is highest under Low-Income Priority, No Choice. For a given choice rule, equivalent variation per assigned unit is highest under Low-Income Priority and lowest under High-Income Priority, though this is largely offset by differences in rent payments from tenants. Due to these similar patterns, the naive decision rule also leads Equal Priority, Choose One to be the preferred mechanism under a low value degree of inequality aversion, while Low-Income Priority, No Choice is preferred with a very high degree of inequality aversion. High-Income Priority, No Choice is the worst of both worlds; it generates low cost-adjusted EV and still only houses a small fraction of extremely high-need applicants. At any degree of inequality aversion, one can find a better allocation mechanism for the CHA.

These findings suggest that the main results of the paper are not entirely driven by the structure placed on decisions and beliefs. Nevertheless, this robustness check is a limited exploration of all possible alternative structures. In particular, under both the naive and sophisticated decision rules, applicants understand the approximate differences in expected waiting time across developments as well as how the position information provided at the final choice stage translates into the distribution of continued waiting times. Furthermore, they share the *same* beliefs; differences in decisions are attributed entirely to differences in preferences. If applicants differ in their waiting time beliefs or decision rules, the estimates in this paper may overstate the degree of preference heterogeneity. An important direction for future research is to measure how well applicants understand the incentives they face in waitlist mechanisms, and to explore implications for allocation policy.

D Counterfactuals: Computational Details

To compute counterfactual equilibria, I draw one sequence of applicant arrivals along with their departure dates, characteristics, and payoffs, and one sequence of apartment vacancies. For the arrival sequence, I first draw a sequence of characteristics of potential applicants from the distribution estimated in Section 5.1, and then draw flow payoffs given those characteristics using the estimates from Specification (3) of the structural model. Apartment vacancies and exogenous departure dates are drawn from the distributions estimated in Section 5.2.

These sequences are used to compute counterfactual allocations under all mechanisms. In computing features of the equilibrium and allocation, the first 20 years are discarded to allow the waiting list to approach steady state. All applicants are eligible for all 13 public housing developments, and all waiting lists remain open during the entire simulation. This abstracts from temporary list closures (which do occur in practice) in order to focus on the long-run effects of choice and priority in steady state.

To compute equilibria of lottery mechanisms allowing choice, I search for a fixed point between applicants' choices and the implied weights $\{w_{j,C}(\psi_\varphi(y_i))\}_{C\in\mathcal{C}_\varphi}^{j=1,\ldots,J}$. The algorithm works as follows. Iteration q begins with a vector of proposed weights $w^{(q)}$. The following steps then occur:

1. Each applicant's optimal choice is calculated when the applicant believes weights are given by $w^{(q)}$.

2. The waiting list is run, yielding predicted weights $w^{(q)\prime}$ with distance $D^{(q)} = \|w^{(q)\prime} - w^{(q)}\|$. To calculate the predicted weights, I consider the experience of one additional applicant on each possible application date in the simulation (after the 20-year burn-in period). For each possible choice C, list j, and priority group ψ, I calculate the waiting time for list j, T_j, and the indicator $1\{T_j \leq \min_{k\in C} T_k\}$ for whether a development j apartment would arrive before any other development in C if the applicant arrived on that particular date. The weights are then calculated using the sample analog of equation 16:

$$w_j^{C,(q)\prime}(\psi) \equiv \frac{1}{\rho}\sum_{d\in D} e^{-\rho\, t_{j,d}(\psi)} * 1\left\{t_{j,d}(\psi) = \min_{k\in C} t_{k,d}(\psi)\right\},$$

   where $t_{j,d}(\psi)$ is the waiting time for an applicant in priority group ψ arriving on date d, and D is the set of application dates averaged over in the simulation. Thus, as in section 5.2, expectations are constructed from the empirical distribution of waiting times generated by the simulation. This method has the benefit of fully accounting for transition dynamics, as the number and types of applicants in the queue fluctuate over time. Transition dynamics are particularly difficult to model through steady-state approximations (Agarwal et al., 2019).

3. Weights are updated as a convex combination of the proposed and implied weights:

$$w^{(q+1)} = \lambda^{(q)} w^{(q)\prime} + (1-\lambda^{(q)}) w^{(q)}.$$

The factor λ determines how aggressively the weights are updated. If $\lambda = 1$, then the weights implied by applicant choices ($w^{(q)'}$) are taken as the new proposal. If $\lambda = 0$, the weights are not updated at all. I began with $\lambda^{(0)} = 1$ and lowered it by 50% each time the Euclidean distance between the proposed and implied offer rates was higher than in the previous iteration ($D^{(q+1)} > D^{(q)}$). This algorithm converged quickly, requiring no more than 50 iterations before implied offer rates were less than 0.1% different than proposed rates in every mechanism. While multiple equilibria are theoretically possible under some of the development choice systems considered in the paper, I did not find multiple equilibria for any mechanism when alternative starting values for the proposed weights $w^{(0)}$ were used.

In the full-information allocations, the social planner uses a greedy algorithm to house applicants from the waiting list. When maximizing equivalent variation from assignments, the planner assigns each vacancy to the applicant with the highest value currently on the waiting list. This is not the strictly optimal policy because each applicant has different values for each development; it may be better to save the highest-value applicant for later and house a lower-value one in a currently vacant unit. These results are still useful as a lower bound on the achievable welfare gains under these ideal conditions. The targeting-maximizing allocation also uses a similar greedy algorithm, assigning each vacancy to the applicant with the worst outside option who is willing to accept the unit.

Table 11: Final Choice Balance Tests

	p-value	F-statistic
Panel A: List Position		
Date of Application	0.000	2.21
Date of Final Choice Letter	0.000	2.82
Annual Income	0.817	0.91
Has Labor Income	0.990	0.78
Has Public Assistance Income	0.700	0.95
Lives in Cambridge	0.582	0.98
Works in Cambridge	0.450	1.01
# Household Members	0.000	1.38
# Earners	0.981	0.81
# Adults	0.298	1.05
# Children	0.004	1.29
White Household Head	0.942	0.85
Black Household Head	0.526	0.99
Hispanic Household Head	0.950	0.84
Age of Household Head	0.281	1.06
Male Household Head	0.084	1.14
# Children under Age 10	0.156	1.10
Panel B: Continued Waiting Time		
Date of Application	0.000	3.44
Date of Final Choice Letter	0.000	2.35
Annual Income	0.982	0.80
Has Labor Income	0.988	0.79
Has Public Assistance Income	0.802	0.92
Lives in Cambridge	0.389	1.03
Works in Cambridge	0.489	1.00
# Household Members	0.000	1.40
# Earners	0.984	0.80
# Adults	0.194	1.09
# Children	0.002	1.33
White Household Head	0.870	0.89
Black Household Head	0.893	0.88
Hispanic Household Head	0.888	0.88
Age of Household Head	0.093	1.14
Male Household Head	0.620	0.97
# Children under Age 10	0.113	1.12

Notes: F-statistics and p-values from a joint test of significance for regression coefficients predicting applicant characteristics as a function of relative list lengths at final choice. Panel A measures list length using list position number, while Panel B uses expected continued waiting time. A different applicant characteristic is the dependent variable in each row.

Table 12: Testing for Responsiveness to List Position at Initial Choice

Development	p-value	F-statistic	DF(1)	DF(2)
Panel A: List Positions Only				
Corcoran Park	0.517	0.981	57	1661
East Cambridge	0.033	1.373	59	1660
Jackson Gardens	0.679	0.905	59	1660
Jefferson Park	0.114	1.235	57	1661
Lincoln Way	0.440	1.018	59	1660
Mid Cambridge	0.002	1.647	59	1660
Newtowne Court	0.090	1.266	57	1661
Putnam Gardens	0.458	1.009	57	1661
River Howard Homes	0.327	1.075	59	1660
Roosevelt Low-Rise	0.114	1.236	57	1661
Washington Elms	0.041	1.358	57	1661
Woodrow Wilson	0.084	1.269	59	1660
Roosevelt Mid-Rise	0.494	0.982	30	1144
Panel B: Applicant Covariates				
Corcoran Park	0.635	0.925	57	1641
East Cambridge	0.189	1.163	59	1640
Jackson Gardens	0.728	0.881	59	1640
Jefferson Park	0.175	1.177	57	1641
Lincoln Way	0.646	0.921	59	1640
Mid Cambridge	0.022	1.415	59	1640
Newtowne Court	0.243	1.127	57	1641
Putnam Gardens	0.262	1.115	57	1641
River Howard Homes	0.700	0.895	59	1640
Roosevelt Low-Rise	0.142	1.206	57	1641
Washington Elms	0.083	1.276	57	1641
Woodrow Wilson	0.101	1.246	59	1640
Roosevelt Mid-Rise	0.408	1.040	30	1125

Notes: F-statistics and p-values from tests for whether list positions predict applicants' choices at initial application. The sample is applicants in the structural estimation sample. For each development, the probability that each applicant chose that development initially is predicted as a function of the length of each list in its choice set. The F-statistic jointly tests for the significance of all coefficients on list position. Panel B adds controls for household income, race/ethnicity, and neighborhood of current residence if the household already lives in Cambridge.

Table 13: Inputs to Waiting Time Simulation

Parameter	Value
Panel A: Apartment Vacancies	
Annual Vacancy Rate per Unit	0.10
Annual Vacancy Rate Total	108
Panel B: Applicant Arrivals and Departures	
Daily Applicant Arrival Rate	0.945
Annual Applicant Arrival Rate	345
Instant Departure Probability	0.243
Annual Departure Rate	0.245
Panel C: Final Choice Model	
List Position Coefficient	-0.019
Fixed Effects	
Corcoran Park	0.347
East Cambridge	-0.130
Jackson Gardens	0.292
Jefferson Park	-0.434
Lincoln Way	0.690
Mid Cambridge	0.265
Newtowne Court	0.073
Putnam Gardens	-0.299
River Howard Homes	0.000
Roosevelt Low-Rise	-0.604
Washington Elms	-0.321
Woodrow Wilson	-0.260
Roosevelt Mid-Rise	-0.876

Table 14: Coefficient Estimates Predicting Probability in CHA Dataset

	Point Estimate	90% Confidence Interval
Income $0-$8,000	2.13	[0.71 , 16.83]
Income $8,000-$16,000	1.64	[0.45 , 13.05]
Income $16,000-$32,000	0.64	[-0.14 , 6.67]
Income $32,000-$48,000	-4.98	[-8.6 , -1.29]
Income Above $48,000	-6.15	[-14.69 , -2.19]
African American Household Head	4.98	[2.18 , 15.24]
Hispanic Household Head	-0.38	[-1.49 , 6.1]
Household lives in Cambridge	-2.19	[-7.02 , 0.95]

Notes: Coefficient estimates predicting the probability that an eligible household from the American Community Survey was in the CHA dataset. The model uses a probit link function and is estimated by minimum distance. The point estimates use the actual ACS 2010-2014 5 percent sample. The 90 percent confidence intervals are bootstrapped by re-sampling the ACS with replacement and re-running the estimation procedure.

Table 15: Simulated Waiting Times from Initial Application

	Simulation		Data	
Development	Mean	S.D.	Mean	# Obs.
Corcoran Park	2.74	1.20	3.05	45
East Cambridge	5.11	1.98	3.52	11
Jackson Gardens	6.14	1.84	3.75	9
Jefferson Park	0.98	1.11	2.16	62
Lincoln Way	3.90	2.19	3.72	2
Mid Cambridge	5.35	2.08	3.52	11
Newtowne Court	2.07	0.95	2.33	95
Putnam Gardens	3.25	1.02	2.98	36
River Howard Homes	6.18	2.17	3.52	11
Roosevelt Low-Rise	2.22	0.87	3.55	21
Washington Elms	2.30	1.39	2.92	26
Woodrow Wilson	4.13	1.69	1.98	2
Roosevelt Mid-Rise	5.03	1.85	1.58	18

Notes: Realized waiting times are averaged across all applicants housed in each development during the simulation.

Table 16: Parameter Estimates, Full

	Baseline Specification		Richer Observed Heterogeneity		Unobserved Taste for Size and Location	
	(1)		(2)		(3)	
Annual Discount Rate	0.973	(0.015)	0.975	(0.013)	0.972	(0.015)
Development 1 Fixed Effect	-0.185	(0.026)	-0.089	(0.03)	-0.110	(0.035)
Development 2 Fixed Effect	-0.794	(0.03)	-0.180	(0.06)	-0.203	(0.068)
Development 3 Fixed Effect	0.034	(0.039)	-0.001	(0.042)	0.056	(0.034)
Development 4 Fixed Effect	-0.052	(0.021)	0.022	(0.018)	-0.007	(0.044)
Development 5 Fixed Effect	0.032	(0.019)	0.004	(0.039)	0.068	(0.026)
Development 6 Fixed Effect	0.000	(0.045)	-0.026	(0.039)	0.003	(0.029)
Development 7 Fixed Effect	-0.418	(0.101)	-0.139	(0.021)	-0.150	(0.035)
Development 8 Fixed Effect	-0.171	(0.06)	-0.169	(0.042)	-0.158	(0.045)
Development 9 Fixed Effect	-0.009	(0.029)	0.085	(0.029)	0.054	(0.037)
Development 10 Fixed Effect	0.000	(0.016)	-0.025	(0.016)	0.000	(0.045)
Development 11 Fixed Effect	-0.488	(0.123)	-0.545	(0.155)	-0.347	(0.052)
Development 12 Fixed Effect	-0.408	(0.107)	-0.246	(0.063)	-0.207	(0.064)
Development 13 Fixed Effect	0.046	(0.025)	0.043	(0.025)	0.043	(0.052)
S.D. Development Fixed Effects	0.263		0.167		0.131	
Panel A: Value of Assistance						
Head Is Black	0.689	(0.047)	0.379	(0.046)	0.395	(0.044)
Head Is Hispanic	-0.011	(0.1)	-0.163	(0.068)	-0.194	(0.112)
Lives In Cambridge	0.487	(0.048)	0.210	(0.032)	0.201	(0.04)
Youngest Member < 10 Years			0.293	(0.038)	0.293	(0.04)
3 Bedroom Household			0.079	(0.054)	0.080	(0.031)
Household Income < $20,000			0.379	(0.05)	0.378	(0.095)
Log Of Observed Income	0.011	(0.047)	0.234	(0.047)	0.231	(0.053)
Log Of Observed And Unobserved Income	-1.000	--	-1.000	--	-1.000	--
Scale of R.E. Unknown Income ($10,000)	1.931	(0.09)	1.716	(0.091)	1.719	(0.127)
Panel B: Match Values						
Applicant and Development Same Neighborhood	0.020	(0.084)	-0.216	(0.047)	-0.249	(0.068)
Applicant Head Is Hispanic * Development Size			0.128	(0.04)	0.126	(0.037)
Youngest Member < 10 Years * Development Size			0.039	(0.015)	0.031	(0.018)
Household Income < $20,000 * Development Size			-0.025	(0.021)	-0.022	(0.02)
S.D. Unobserved Taste For Development Size					0.031	(0.013)
S.D. Unobserved Taste for North Cambridge					0.031	(0.012)
S.D. Unobserved Taste for East Cambridge					0.005	(0.004)
S.D. Idiosyncratic Shock	0.152	(0.01)	0.124	(0.007)	0.124	(0.01)

Table 17: Effects of Alternative Development Choice Systems, Naive Decision Rule

	Choose One (1)	Choose Any Subset (2)	Choose All or One (3)	Choose Neighborhood (4)	Choose All or Neighborhood (5)	No Choice (6)
Panel A: Welfare Gain and Cost of Allocation						
Equivalent Variation ($)	17,059	17,058	17,063	13,318	13,324	11,774
Cost per Unit ($)	18,220	18,226	18,221	18,888	18,888	19,196
Equivalent Variation per $ Cost to Gvt.	0.94	0.94	0.94	0.71	0.71	0.61
Panel B: Targeting						
Observed Income ($)	19,267	19,246	19,265	17,040	17,040	16,013
Observed and Unobserved Income ($)	13,347	13,326	13,341	11,038	11,035	9,826
% Extremely High-Need	31.9%	32.0%	31.9%	37.5%	37.5%	41.2%
Panel C: Match Quality						
% Assigned Top Choice	83.7%	83.2%	83.6%	22.8%	22.8%	11.9%
% Assigned Top 3	99.2%	99.1%	99.2%	50.9%	51.0%	31.4%
Panel D: Characteristics of Housed Applicants						
Waiting Time (days)	1136	1137	1135	926	926	824
% Black	46.5%	46.5%	46.5%	48.7%	48.7%	49.5%
% Hispanic	24.4%	24.4%	24.4%	20.2%	20.2%	19.1%
From Cambridge	58.4%	58.4%	58.4%	58.0%	58.0%	59.1%

Notes: Counterfactual estimates are obtained identically to those in Table 9, but using the preference estimates based on a naive development choice rule.

Table 18: Alternative Choice and Priority Systems, Naive Decision Rule

	Common Choice and Priority Systems						Full Information	
	Low-Income Priority		High-Income Priority		Equal Priority		Equivalent Variation Maximizing	Targeting Maximizing
	Choose One	No Choice	Choose One	No Choice	Choose One	No Choice		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Welfare Gain and Cost of Allocation								
Equivalent Variation ($)	18,911	12,671	14,627	11,033	17,059	11,774	24,122	13,799
Cost per Unit ($)	20,558	21,468	15,847	17,194	18,220	19,196	19,518	20,251
Equivalent Variation per $ Cost to Gvt.	0.92	0.59	0.92	0.64	0.94	0.61	1.24	0.68
Panel B: Targeting								
Observed Income ($)	11,473	8,439	27,178	22,688	19,267	16,013	14,938	12,496
Observed and Unobserved Income ($)	10,135	7,756	16,441	11,690	13,347	9,826	9,713	4,335
% Extremely High-Need	43.1%	52.9%	20.6%	31.3%	31.9%	41.2%	46.0%	72.6%
Panel C: Match Quality								
% Assigned Top Choice	77.2%	9.5%	82.7%	13.8%	83.7%	11.9%	43.6%	7.9%
% Assigned Top 3	96.6%	27.5%	96.1%	34.5%	99.2%	31.4%	73.7%	24.3%
Panel D: Characteristics of Housed Applicants								
Waiting Time (days)	735	236	735	491	1136	824	113	69
% Black	47.8%	51.5%	45.3%	47.8%	46.5%	49.5%	75.5%	38.9%
% Hispanic	24.8%	20.7%	21.8%	17.8%	24.4%	19.1%	16.0%	20.7%
From Cambridge	62.8%	64.3%	53.7%	55.0%	58.4%	59.1%	54.8%	64.4%

Notes: Counterfactual estimates are obtained identically to those in Table 10, but using the preference estimates based on a naive development choice rule.

Table 19: Effects of Alternative Development Choice Systems under Low-Income Priority

	Choose One	Choose Any Subset	Choose All or One	Choose Neighborhood	Choose All or Neighborhood	No Choice
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Welfare Gain and Cost of Allocation						
Equivalent Variation ($)	15,444	15,355	15,443	11,538	11,538	11,274
Cost per Unit ($)	19,948	19,937	19,949	18,969	18,969	21,240
Equivalent Variation per $ Cost to Gvt.	0.77	0.77	0.77	0.61	0.61	0.53
Panel B: Targeting						
Observed Income ($)	13,505	13,542	13,505	16,769	16,769	9,200
Observed and Unobserved Income ($)	8,202	8,234	8,203	8,091	8,096	7,183
% Extremely High-Need	51.9%	51.7%	51.9%	47.7%	47.7%	56.6%
Panel C: Match Quality						
% Assigned Top Choice	41.6%	40.3%	41.6%	11.2%	11.2%	7.0%
% Assigned Top 3	63.8%	63.4%	63.8%	29.5%	29.6%	21.4%
Panel D: Characteristics of Housed Applicants						
Waiting Time (days)	850	872	851	705	706	289
% Black	44.7%	44.7%	44.7%	42.5%	42.5%	48.0%
% Hispanic	21.6%	21.5%	21.6%	19.0%	19.0%	19.5%
From Cambridge	65.1%	65.1%	65.1%	64.0%	64.0%	67.8%

Notes: Statistics averaged across assigned apartments in each counterfactual simulation. Dollar amounts are annual. Cost per Unit is calculated based on market rate rental prices in Cambridge, MA during the sample period. Equivalent Variation is the equivalent cash transfer outside of public housing that would generate the same welfare change for a housed applicant as their assignment. Extremely High-Need applicants are at the minimum consumption level of $10/day outside of public housing. Low-Income Priority first offers vacant apartments to applicants with incomes below $15,000.

Table 20: Effects of Alternative Development Choice Systems under High-Income Priority

	Choose One (1)	Choose Any Subset (2)	Choose All or One (3)	Choose Neighborhood (4)	Choose All or Neighborhood (5)	No Choice (6)
Panel A: Welfare Gain and Cost of Allocation						
Equivalent Variation ($)	12,944	12,913	12,944	10,345	10,345	9,651
Cost per Unit ($)	17,302	17,302	17,302	18,444	18,444	16,914
Equivalent Variation per $ Cost to Gvt.	0.75	0.75	0.75	0.56	0.56	0.57
Panel B: Targeting						
Observed Income ($)	22,326	22,325	22,326	18,521	18,521	23,619
Observed and Unobserved Income ($)	10,711	10,707	10,711	9,241	9,241	9,473
% Extremely High-Need	40.0%	39.9%	40.0%	44.6%	44.6%	38.2%
Panel C: Match Quality						
% Assigned Top Choice	45.1%	44.6%	45.1%	13.5%	13.5%	8.7%
% Assigned Top 3	62.8%	62.7%	62.8%	31.0%	31.0%	23.4%
Panel D: Characteristics of Housed Applicants						
Waiting Time (days)	1075	1085	1075	915	915	443
% Black	40.0%	40.0%	40.0%	41.7%	41.7%	38.2%
% Hispanic	18.8%	18.9%	18.8%	18.6%	18.6%	16.4%
From Cambridge	62.4%	62.4%	62.4%	64.6%	64.6%	62.1%

Notes: Statistics averaged across assigned apartments in each counterfactual simulation. Dollar amounts are annual. Cost per Unit is calculated based on market rate rental prices in Cambridge, MA during the sample period. Equivalent Variation is the equivalent cash transfer outside of public housing that would generate the same welfare change for a housed applicant as their assignment. Extremely High-Need applicants are at the minimum consumption level of $10/day outside of public housing. High-Income Priority first offers vacant apartments to applicants with incomes above $15,000.

Figure 5: Locations of Cambridge Family Public Housing Developments

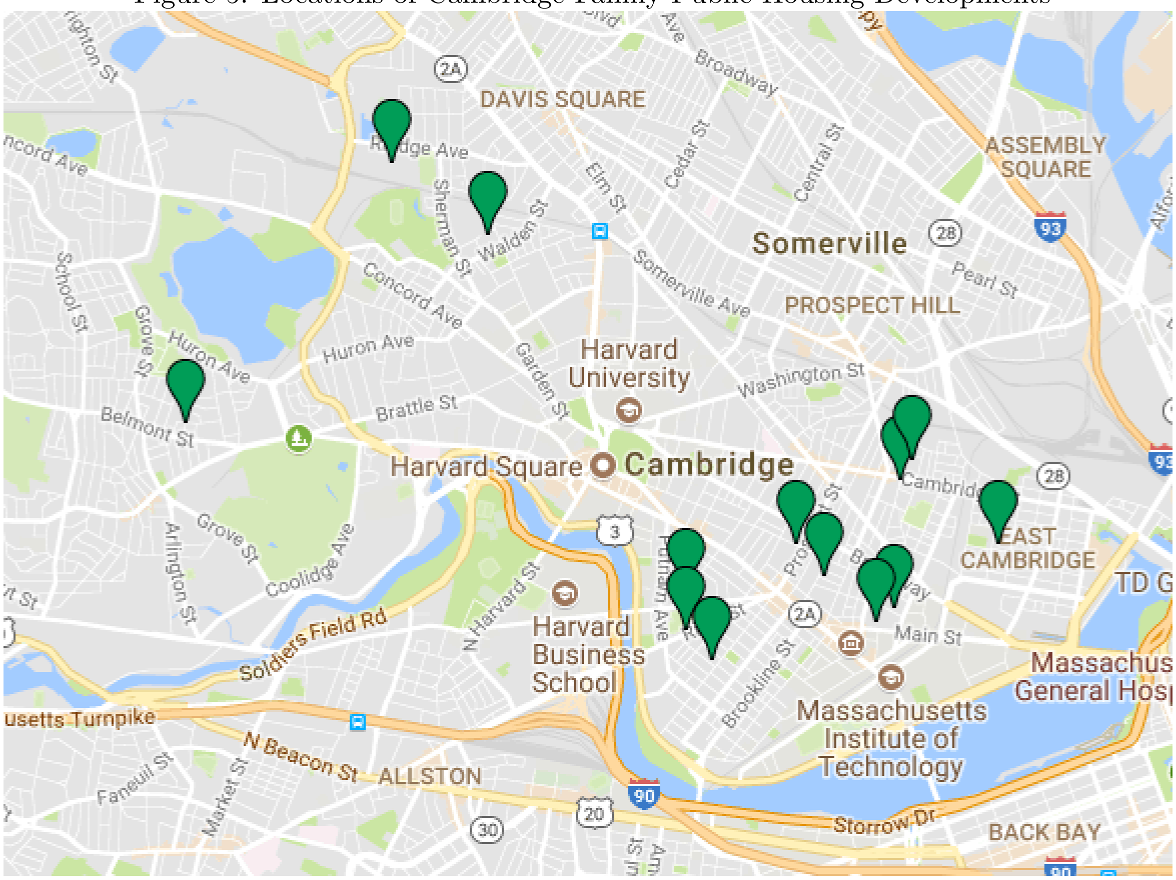
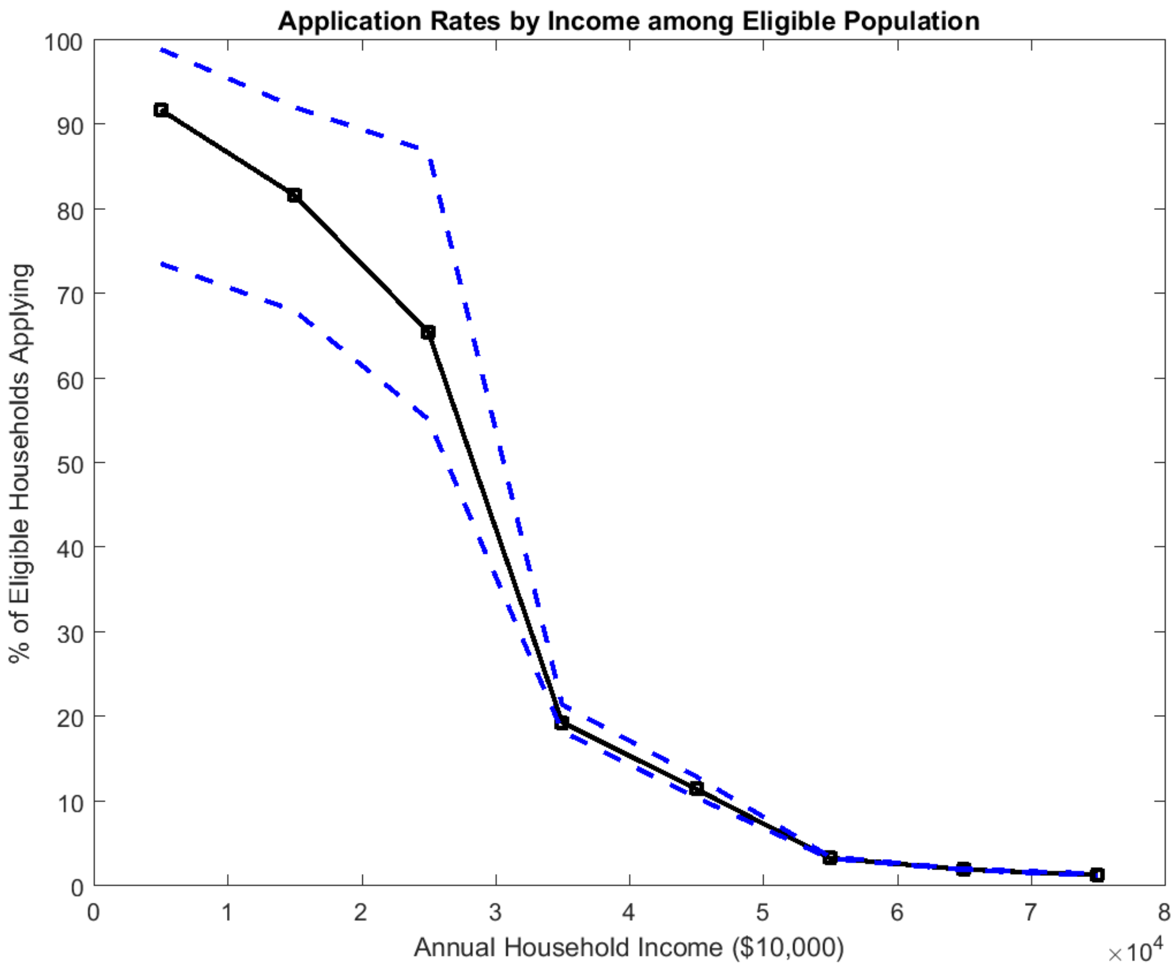


Figure 6: Application Rates by Income



Notes: The estimated fraction of eligible households that applied for Family Public Housing in Cambridge between 2010 and 2014, by $10,000 income groups. For each group, the number of applicants is divided by the number of eligible households as estimated in Section 6.1. The dotted lines give point-wise 90 percent confidence bands obtained from a bootstrap that re-samples the set of eligible ACS households with replacement.

Figure 7: Welfare Under Alternative Development Choice Systems, Low-Income Priority

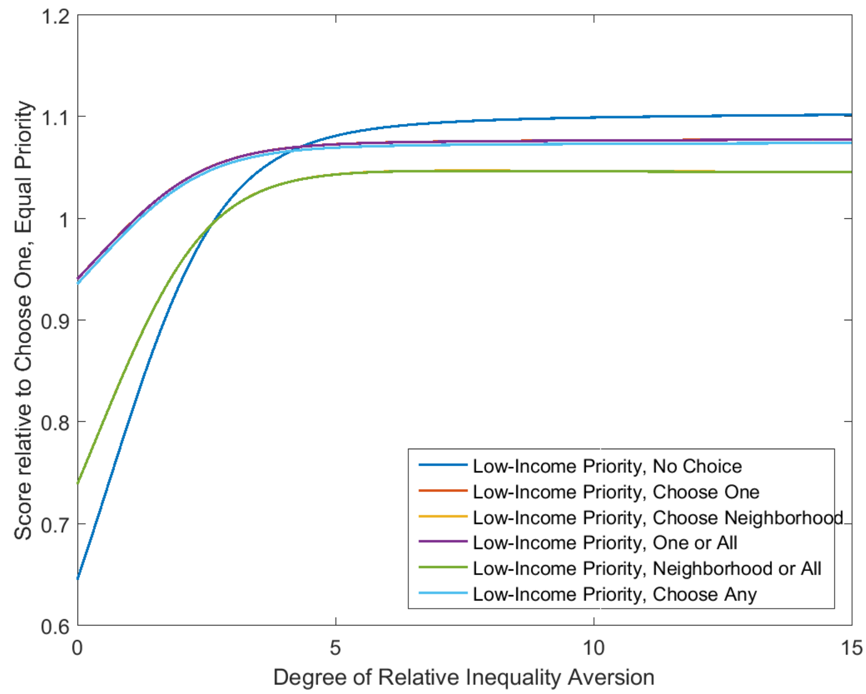
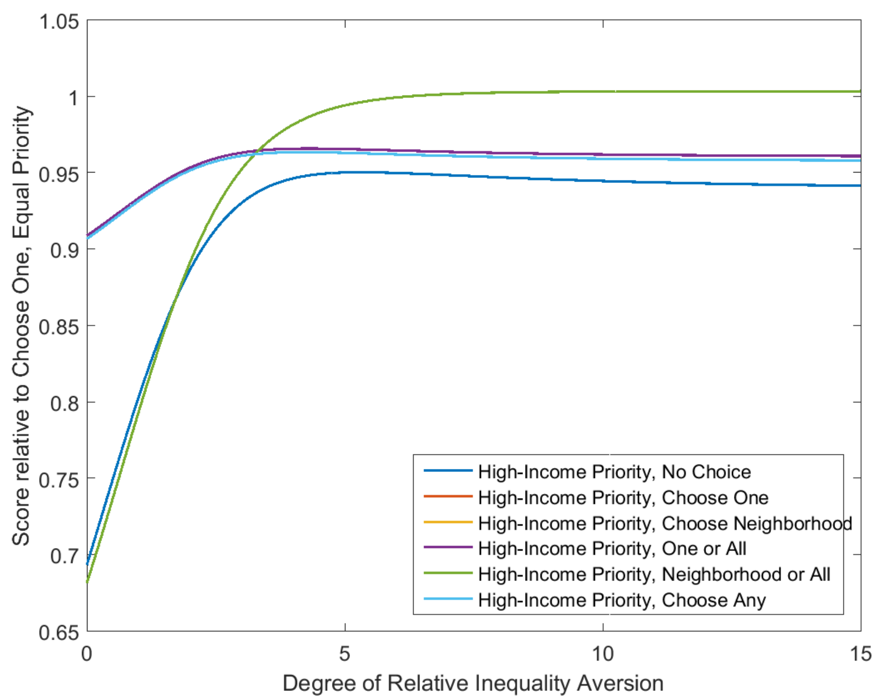
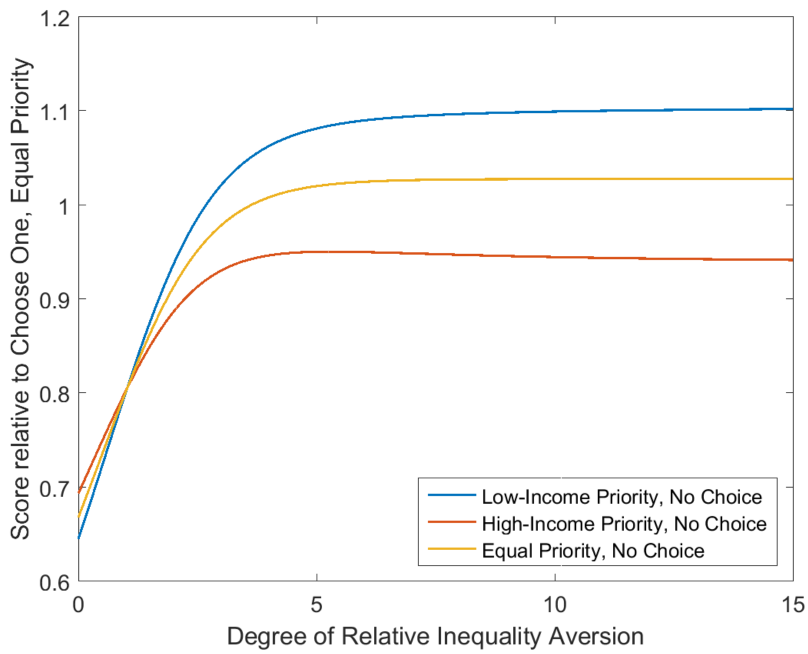


Figure 8: Welfare Under Alternative Development Choice Systems, High-Income Priority



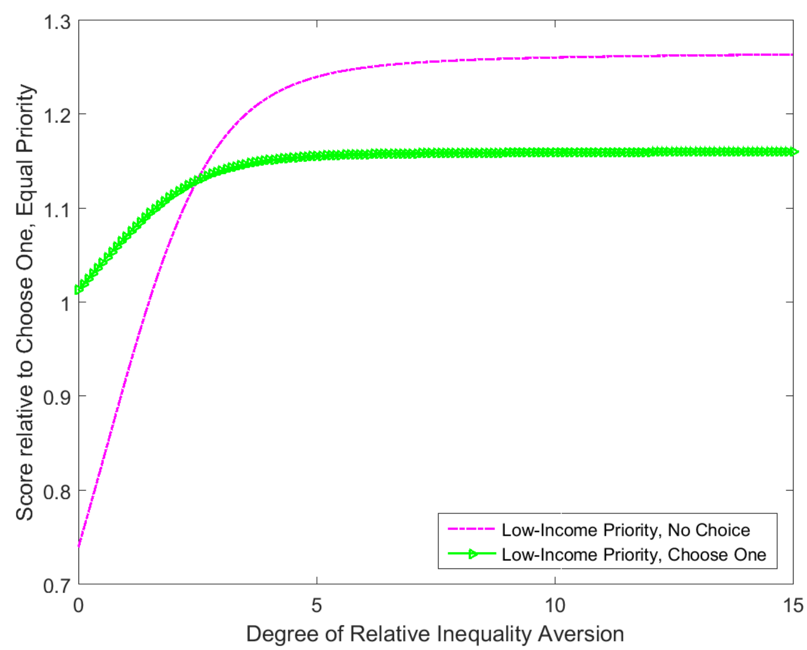
Notes: Figures 7 and 8 compare cost-adjusted welfare gains produced by development choice systems used in practice, with priority for households with incomes below and above \$15,000, respectively. Welfare gains are normalized by the value for Equal Priority, Choose One.

Figure 9: Welfare Under Alternative Priority Systems with No Choice



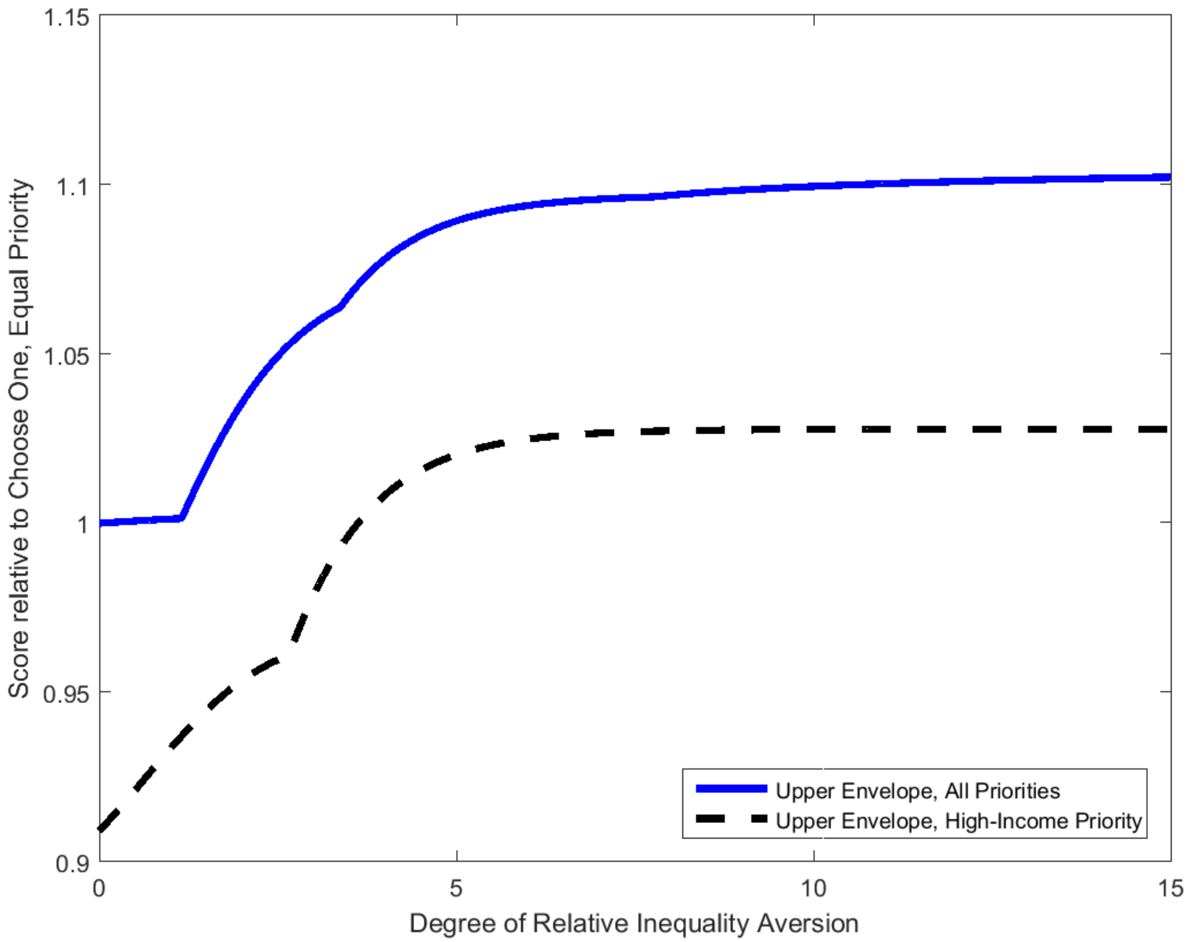
Notes: Cost-adjusted welfare gains produced by alternative priority systems under No Choice. Low-Income and High-Income Priority prioritize households with annual incomes below and above $15,000, respectively. Welfare gains are normalized by the value for Equal Priority, Choose One.

Figure 10: Preferred Mechanisms by Degree of Inequality Aversion, without Cost Adjustment



Notes: Welfare effects of mechanisms which perform best for some degree of inequality aversion without adjusting for rent payment differences across households of different incomes. Welfare gains are normalized by the value for Equal Priority, Choose One.

Figure 11: High-Income Priority Mechanisms Dominated



Notes: Upper Envelope, All Priorities plots the highest cost-adjusted welfare gain among all choice and priority rules considered for each degree of inequality aversion. Upper Envelope, High-Income Priorities plots the analogous quantity among all mechanisms with High-Income Priority. Welfare gains are normalized by the value for Equal Priority, Choose One.