Time Use and the Efficiency of Heterogeneous Markups[*]

Brian C. Albrecht[†]   Thomas Phelan[‡]   Nick Pretnar[§]

January 3, 2025

Abstract

What are the welfare implications of markup heterogeneity across firms? In standard monopolistic competition models, markup heterogeneity implies inefficiency even in the presence of free entry. We enrich the standard model so that preferences depend non-separably on off-market time and show how this changes the equilibrium and efficient distributions of productivity and markups. With constant elasticity of substitution across varieties, firm selection is inefficiently lax when off-market time and market goods are complements and inefficiently strict when they are substitutes. However, when off-market time and market goods are perfect complements, markups differ across firms and yet the equilibrium allocation is efficient.

Keywords: monopolistic competition, markups, efficiency, time use, elasticity of substitution, selection, heterogeneous firms.

JEL codes: D1, D4, D6, L1

1 Introduction

Studies of markups find large heterogeneity across firms.[1] In standard models of monopolistic competition with endogenous entry, such as Dhingra and Morrow (2019), consumers care solely about varieties of consumption goods and heterogeneous markups imply misallocation. In this paper, we study markups and efficiency in an environment in which off-market time enters the consumer's preferences non-separably in the spirit of Becker (1965). We show how time use alters the distribution of markups and that heterogeneous markups can be consistent with allocative efficiency.

In our model, consumers have preferences over varieties of consumption "experiences" that are produced using both market goods and off-market time. We think of these experiences as time-consuming activities such as attending a play or playing video games, in which the market price of the good is not the total price of the associated experience. For example, attending a play requires both a ticket (a market purchase) and the attendee's time, which could have been devoted to other experiences or to earning labor income. Because consumers can respond to price changes by adjusting both off-market time and market purchases, demand elasticities (and hence markups) depend on the degree to which off-market time can substitute for market goods. In this way, non-separable time use provides a source of variation in markups beyond that generated by variable elasticities of substitution across experiences.

We first decompose the role of non-separable time use on demand elasticities into two opposing effects. First, the fact that time use serves as an input into final consumption implies that the total price of each experience is relatively less elastic with respect to the market price. For instance, if the price of a theatre ticket doubles, the total price to the consumer of attending the play less than doubles, because their opportunity cost is unchanged. For a given elasticity of substitution across varieties, non-separable time use therefore reduces the price elasticity of each experience. Second, as the price of a market good rises, consumers can increase the share of time use in final consumption, which serves to increase the price elasticity of (market) demand and implies that the net impact of the above two effects is theoretically ambiguous. With CES preferences over experiences, markups rise with firm productivity (i.e., fall as firms' marginal costs fall) and firm size when off-market time and market goods are complements but may either rise or fall with productivity in the case of substitutes. In the special cases of unit elastic substitutability or additive separability between off-market time and market purchases, markups are constant across firms.

Because time use affects demand elasticities and firm profits, it alters the incen-

---

[1]See, e.g., Epifani and Gancia (2011), De Loecker et al. (2020), Peters (2020), and Edmond et al. (2023).

tives for firms to enter and operate. Following Dhingra and Morrow (2019), we model firm entry as a two-step process. First, ex-ante identical firms must pay a fixed cost (entry cost) in order to draw a parameter specifying their marginal cost of production. Second, upon learning their marginal cost, firms must pay a second fixed cost (operating cost) in order to produce. There are therefore two quantities related to firm entry: the mass of firms that draw a productivity parameter, and the productivity of the marginal firm, which we refer to as firm "selection." We then study whether the equilibrium values of these quantities are higher or lower than those chosen by a utilitarian planner.

To facilitate this analysis, we first show that equilibrium allocations solve a certain optimization problem, in which the objective is proportional to the aggregate revenues of all firms and the constraint incorporates the role of time use. We use this insight to show that the analysis of firm selection amounts to comparing the elasticities of two quantities: the first representing the average *social* value of all varieties, and the second representing the average *private* value of all varieties (i.e., profits). Further, with CES preferences over experiences, we show that firm selection depends solely on the elasticity of substitution between time and market consumption: when market goods and time are substitutes, too few entering firms survive, while when they are complements, too many do. More precisely, in equilibrium, the marginal cost of the marginal firm is (weakly) inefficiently high when off-market time and market purchases are complements and inefficiently low when they are substitutes. However, in the special cases in which off-market time and market purchases are perfect complements or exhibit unit elasticity of substitution, the marginal cost of the marginal firm coincides with the efficient value.

The intuition for the above selection results is as follows. When off-market time and market consumption are perfect complements, they are consumed in fixed proportions independently of prices, and so the model is essentially identical to a model with no time use (and relabeled marginal costs), where efficiency is known to obtain with CES preferences.[2] Outside of this extreme case, when time and consumption are complements, consumers cannot easily substitute time for market consumption as costs rise, and so profits are less elastic than social surplus. Conversely, when market goods and time are substitutes, the reverse is true because consumers can more easily substitute time for market purchases, and effectively use their time to compete with firms. This ensures that profits fall more rapidly than social surplus as costs rise and leads to an inefficiently large number of firms that choose not to operate.

We then turn to the aggregate resources devoted to firm entry, and show that the equilibrium allocation always devotes (weakly) too few resources to firm entry and

---

[2] See, e.g., Dhingra and Morrow (2019).

product creation, regardless of the structure of the consumption technology. Consumers in this economy do not internalize the fact that their decisions to allocate labor between off-market time and market production affect firms' profits and the incentives to enter. This is an additional margin of adjustment that is not present in models without off-market time. However, the sole exception in terms of efficiency occurs in the case of perfect complements, in which the ratio of off-market time to market purchases is independent of prices and marginal costs. In this case, the mass of firms also coincides with the efficient value, and so the equilibrium allocation in this case is first-best.

To interpret this last result and relate to the literature, we distinguish between two different notions of markups. The first is the usual definition of a markup ratio: the price of a good divided by its marginal cost. However, because market goods are only one input into consumption experiences, we contrast the usual markup ratio with an alternative, "holistic" notion: the ratio of the total price of an experience to its total marginal cost, where both account for the value of the consumer's time. Equipped with this nomenclature, we believe that the efficiency of the perfect complements case becomes intuitive, because this is the sole situation in which holistic markups (but not the usual markups) are constant across firms.

The results in this paper have several important implications. First, our result that efficiency attains when off-market time and market goods are perfect complements in the consumption technology challenges the idea that variable markups always imply welfare losses.[3] This is potentially important given that estimates of markup dispersion suggest an increase over the last 40 years.[4] Second, the fact that holistic markups are always less than the usual markups implies that an exclusive focus on the latter may overstate both the welfare implications of markups and market concentration.

Related literature. The seminal contribution to the theory of optimal product variety is Dixit and Stiglitz (1977), who consider an environment with homogeneous firms and endogenous entry and show that the efficiency of the equilibrium depends on whether preferences over varieties exhibit constant or variable elasticities of substitution. The literature emanating from Dixit and Stiglitz (1977) is vast and so we highlight only closely related developments.[5] Zhelobodko et al. (2012) provide further insight into the variable elasticity of substitution (VES) equilibrium with homogeneous firms and derive general comparative statics. Dhingra and Morrow (2019) extend the analysis of optimal product variety to models with heterogeneous firms, while Behrens et al. (2020) allow for different elasticities between sectors and quantify

---

[3]In models such as Peters (2020) and Edmond et al. (2023), markup dispersion creates misallocation and ultimately reduces welfare relative to a benchmark with no dispersion.

[4]See, e.g., Figure 3 of De Loecker et al. (2020) and Figure 2 of Flynn et al. (2019).

[5]See the special issue in Etro (2017) for further discussion and historical context.

the welfare losses associated with markups. In these papers time use is inelastic and CES preferences are both necessary and sufficient for markups to be constant and for allocations to be efficient when marginal costs differ across firms. Our results are reminiscent of Parenti et al. (2017), who enrich the problem of the consumer to incorporate uncertainty over love-for-variety, and also show that variable markups can be consistent with efficiency.

Our modeling of the consumer problem reflects the fundamental idea first discussed in Becker (1965) and analyzed in depth in Ghez and Becker (1974) that in order to enjoy consumption, consumers must allocate time toward doing so. This is the basis of the "consumption technology" we implement: just like a production technology it takes in intermediate inputs (i.e., goods and services bought on the market) and time to generate an output, which we call "final consumption" or a "consumption experience."[6] Our approach is motivated by empirical evidence that time use and market purchases (what we sometimes refer to as "market consumption") are not separable in preferences.[7] The literature has recently begun to consider the implications of incorporating non-separable preferences for market purchases and leisure time for outcomes only indirectly related to consumer time use itself. For example, Boerma and Karabarbounis (2021) and Pretnar (2024) use models with consumption and time-use non-separabilities to measure welfare inequality, Bridgman et al. (2018b) and Bednar and Pretnar (2024) study how accounting for time use affects implications for structural change, and Bridgman (2016) studies how in-home productivities for different types of market products have changed over time. We add to these findings by studying the welfare implications of the Becker (1965) model in an economy with monopolistically competitive firms.

Our incorporation of non-separable off-market time into the consumer problem builds on models of monopolistic competition which also allow for elastic labor supply. The literature on monopolistic competition has considered elastic labor, but to the best of our knowledge it has assumed separability from consumption. Bilbiie et al. (2012) explore how the allocation of labor across sectors and the number of products and producers varies over the business cycle. Bilbiie et al. (2019) consider the role that elastic factors of production (labor and capital) play in amplifying distortions on a dynamic path along which firms may enter and exit. Boar and Midrigan (2019)

---

[6]In the literature some refer to this process as "home production," though the term, which seems to have originated with Gronau (1977), tends to have a more narrow interpretation than what Becker (1965) originally theorized. The Becker (1965) approach can be seen in the models of Greenwood and Hercowitz (1991), Gomme et al. (2001), Greenwood et al. (2005), Ngai and Pissarides (2008), and Bridgman et al. (2018a) where home technologies are assumed to take in both time and market select goods (i.e., durables, physical goods, etc.) to yield a home production output.

[7]See, e.g., Aguiar and Hurst (2005), Aguiar and Hurst (2007), Fang et al. (2022) and Pretnar (2024).

allow for additively separable consumption and time, but focus on how markup distortions redistribute income from laborers to entrepreneurs. De Loecker et al. (2021) also allow for separable consumption and off-market time but only include leisure in preferences to quantify how changes to the structure of competition affect the labor supply. Finally, Edmond et al. (2023) disentangle the degree to which markups, misallocation of factors of production, and inefficient entry contribute to welfare costs, and quantify the welfare effects of markups. We add to this literature by considering how the allocation of off-market time affects general equilibrium outcomes in a monopolistically competitive setting.

2 General model

In this section we describe a general model environment where preferences and the consumption technology are not explicitly parameterized. The goal is to characterize both equilibrium and efficient allocations when time use is elastic and off-market time and market purchases are non-separable in the consumer's problem.[8]

The setup is as follows. First, consumers buy varieties of goods and services on the market. These varieties are produced by monopolistically competitive firms using a linear production technology where labor is the only input. Next, consumers use all of the varieties they purchase on the market in various off-market consumption activities. In doing so they are allocating their resources to an intermediate production step which involves using two inputs — market varieties and off-market time — in order to produce a final consumption good or experience. Consumers ultimately derive utility from this experience, produced by combining time and market goods in the consumption technology. We think of these final consumption goods as experiences that cannot be bought and sold, despite the fact that the inputs used to make these experiences (i.e., market goods and services, as well as time) are tradable. Becker (1965) referred to "attending a play" as one of these experiences. While tickets to the play can be bought and sold on the market, a ticket (the market good/service) will yield the consumer no utility unless it is actually used for attendance (as opposed to just sitting unused on one's desk at home). Attendance, though, requires a consumer's time as well. Video gaming is a more modern example of a consumption experience: the model of Aguiar et al. (2021) relates consumption expenditure to time allocation decisions similar to the model we outline. To derive utility from a video-gaming experience, consumers must first buy a game (a software service) and then spend

---

[8]Bilbiie et al. (2012) and Bilbiie et al. (2019) feature variable labor supply and leisure that are separable from market consumption, while in Zhelobodko et al. (2012) and Dhingra and Morrow (2019) time use is inelastic and does not enter into an agent's utility function.

their time playing that game. Simply owning the game (or subscribing to a service that provides access to the game) without actually playing it provides no utility.

The aforementioned examples suggest a model in which consumer demand for market goods depends on both their marginal utility for the final consumption experience and the marginal product of market goods and time in the consumption technology. However, firms have direct control over the price of only one of the consumer inputs (market goods) but not the value of a consumer's off-market time nor the unit-value (or "price") of the final consumption experience. This latter value is a weighted sum of the price of market varieties and the value of a consumer's off-market time. As the following formal model shows, this means that firm pricing rules for market goods will depend on the degree to which such goods are complementary or substitutable with off-market time. The structure of the consumption technology will thus determine the degree to which firms can price above their marginal cost and therefore ultimately which firms choose to operate.

2.1   Consumers

There exists a unit mass of identical consumers who have preferences over a continuum of final consumption *experiences*, each of which is indexed by $i \in \mathcal{I} \subseteq \mathbb{R}$:

$$\mathcal{U}(\mathbf{c}) = \int_{\mathcal{I}} u(c_i) \mathrm{d}i. \tag{2.1}$$

We denote consumption experiences by c_i, where $\mathbf{c} \subset \mathbb{R}_+$ denotes the set, indexed by the set $\mathcal{I}$, of all consumption experiences.[9] We assume that $\mathcal{I} \subset \mathbb{R}_+$ is compact. As in Dhingra and Morrow (2019), we make the following regularity assumptions on u.

Assumption 2.1 (Regularity of u). *The sub-utility function u strictly increasing, strictly concave, at least four-times differentiable, and such that $u(0) = 0$.*

A consumption experience is defined as the act of combining goods and services purchased on the market with off-market time in order to generate final utility.[10] Consumers produce these experiences using a "production function" h that takes market consumption, q_i, and off-market time, n_i, as inputs,

$$c_i := h(q_i, n_i). \tag{2.2}$$

We emphasize that n_i represents the time devoted to the consumption of the ith

---

[9]Bold faced fonts will denote sets, so that $c_i \in \mathbf{c}$.

[10]Our *experiences* are analogous to what Becker (1965), Aguiar et al. (2012), Aguiar and Hurst (2013), and Aguiar and Hurst (2016) refer to as "commodities."

variety q_i, and not a distinct good produced by the household.[11] We will later show in our parametric exercises how the functional form of h (and specifically, whether q_i and n_i are complements or substitutes) affects both the efficient allocations and the distribution of markups. In this section, we adopt the following general assumptions.

Assumption 2.2 (Regularity of h)**.** *The function h is strictly increasing, strictly concave, four-times differentiable, and satisfies the Inada conditions $\lim_{x \to 0^+} h_q(x, n) = \lim_{x \to 0^+} h_n(q, x) = \infty$ for all $q, n \geq 0$.*

Assumption 2.3 (Homogeneity of h)**.** *The function h is homogeneous of degree one.*

Assumption 2.3 is a standard assumption (see, e.g., Aguiar et al. (2012) and Aguiar and Hurst (2016)), and will allow us to interpret our results in terms of the prices of either market consumption or final consumption. Because each consumption experience depends on one and only one market variety, we think of $u(c_i)$ as the utility a consumer derives from the utilization of market variety, q_i. This is similar to how u is defined in Dixit and Stiglitz (1977) and Dhingra and Morrow (2019) but with the added complication that utilization of the market variety requires off-market time. In the video game example, the experience is "playing video games" which requires the game itself, q_i, and time spent playing, n_i. As in standard models of monopolistic competition, consumers exhibit a "love-of-variety," but it is a love of variety for consumption *experiences* rather than solely market consumption.

Resource constraints. Each consumer is endowed with $\overline{T}$ units of time that is split between three types of activities: 1) labor devoted to creating firms, $\mathcal{E}$; 2) labor devoted to the production of market consumption, $\mathcal{L}$; and 3) time devoted to the production of final consumption, $\mathcal{N} := \int_{\mathcal{I}} n_i \, \mathrm{d}i$. The consumer's time-use constraint is the sole resource constraint in the economy and is given by

$$\mathcal{E} + \mathcal{L} + \mathcal{N} \leq \overline{T}. \tag{2.3}$$

Consumers earn labor income, $\mathcal{E} + \mathcal{L}$, on the market, so total expenditure on market goods must satisfy the budget constraint $\int_{\mathcal{I}} p_i q_i \mathrm{d}i = \mathcal{E} + \mathcal{L}$, where p_i is the price of market good/service q_i and we normalize hourly wages to unity. When combined with the constraints (2.2), (2.3), and the objective (2.1), the problem of the consumer is

---

[11]The latter interpretation is closer to Gronau (1977), who sought to expand Becker (1965) to incorporate household production. Our formulation follows Becker (1965), who develops a theory of consumer choice that includes the cost of time on the same footing as the cost of market goods.

to choose market varieties $\mathbf{q}$ and off-market time $\mathbf{n}$ to solve the problem

$$\max_{\mathbf{q},\mathbf{n}} \int_{\mathcal{I}} u(h(q_i, n_i)) \mathrm{d}i \tag{2.4}$$
$$\text{s.t.} \int_{\mathcal{I}} (p_i q_i + n_i) \mathrm{d}i = \overline{T}.$$

If λ is the multiplier on the consumer's budget constraint then the first-order condition with respect to q_i is $u'(h(q_i, n_i))h_q(q_i, n_i) = p_i \lambda$. If $h(q, n) = q$ for all $q, n > 0$, then differentiating the first-order condition with respect to p_i shows that the price elasticity of demand is simply the reciprocal of the elasticity of marginal utility. However, outside of this special case, consumers may choose to adjust their time use in response to a price change, and in so doing affect their marginal utility of market consumption. To see how this affects the price elasticity of demand, we first note that Assumptions 2.2 and 2.3 imply that for some increasing function Γ, the optimal choices of the consumer satisfy $n_i/q_i = \Gamma(p_i)$ for all $i \in \mathcal{I}$.[12] The ith term in the budget constraint in (2.4) may be written as $\psi_i c_i$, where

$$\psi_i := \psi(p_i) := \frac{p_i + \Gamma(p_i)}{h(1, \Gamma(p_i))} = \min_{x > 0} \frac{p_i + x}{h(1, x)} \tag{2.5}$$

can be interpreted as the per-unit of final consumption, inclusive of off-market time.[13] The problem (2.4) can therefore be solved by first considering the problem of maximizing $\int_{\mathcal{I}} u(c_i) \mathrm{d}i$ subject to the constraint $\int_{\mathcal{I}} \psi(p_i) c_i \mathrm{d}i = \overline{T}$, which is identical to a problem without time use except that the price of a unit of the ith variety is not p_i but instead the "holistic price" $\psi(p_i)$.

To derive the price elasticity of demand, we write market consumption as $q = c \times q/c$ and decompose its price elasticity into the price elasticities of c and q/c separately. Differentiating the first-order condition for c_i gives $\epsilon_c(p_i) = \epsilon_\psi(p_i)/\epsilon_{u'}(c(p_i))$, and applying the envelope theorem to (2.5) gives $q(p_i) = c(p_i)\psi'(p_i)$. Combining these two observations then immediately gives the following.

Lemma 2.4 (Demand elasticity)**.** *The price elasticity of demand satisfies*

$$\epsilon_q(p_i) - \frac{1}{\epsilon_{u'}(c(p_i))} = \frac{\epsilon_\psi(p_i) - 1}{\epsilon_{u'}(c(p_i))} + \epsilon_{\psi'}(p_i). \tag{2.6}$$

The expression for the price elasticity of demand in Lemma 2.4 is fundamental to

---

[12]Assumption 2.3 implies that choice of q_i and n_i solve $h_q(1, n_i/q_i)/h_n(1, n_i/q_i) = p_i$, and Assumption 2.2 implies that Γ is well-defined and increasing.

[13]The function ψ_i is analogous to the π_i function introduced on page 422 of Becker (1965). We write ψ instead of π because the latter is now typically reserved for operating profits.

the analysis in this paper and therefore warrants some discussion. As noted above, when final consumption does not depend on time use, the price elasticity of consumption is $\epsilon_{u'}(c(p_i))^{-1}$, and so Lemma 2.4 implies that the difference between the price elasticity of demand and this "usual" term is the sum of two distinct economic effects corresponding to the terms of the right-hand side of (2.6).

First, because it is the minimum of affine functions, the function ψ is concave, and so its elasticity cannot exceed unity.[14] This bound reflects the fact that market consumption is only one input (along with time) into final consumption, which ensures that ψ is inelastic with respect to p_i. For instance, in the theatre example of Becker (1965), if the price of entry were \$30 and the consumer valued the time taken to watch the play at \$100, then a doubling of the ticket price would lead ψ to rise by less than double, from \$130 to \$160.[15] For a given elasticity of marginal utility, the presence of time use always reduces the price elasticity of final consumption and so the first term on the right-hand side of (2.6) is non-negative.

Second, as the price of a variety of market consumption rises, consumers can not only reduce their demand for final consumption, but also adjust the ratio of market consumption to final consumption. By the envelope condition, we have $q(p_i) = c(p_i)\psi'(p_i)$, and so this second effect is captured by the elasticity $\epsilon_{\psi'}$, which is always non-positive by the concavity of ψ, and therefore opposes the first effect above. For example, in the case of video gaming, where utility depends on both gaming equipment purchased and the time spent playing, as the price of gaming equipment rises, consumers can devote more time to each game, and their willingness to substitute time for market consumption is reflected in the elasticity $\epsilon_{\psi'}$.[16] For the functional forms in Section 3, the terms on the right-hand side of (2.6) simplify in a manner that facilitates simple comparative statics for markups.

2.2 Firms

For the modeling of firm entry, technology and behavior, we deliberately follow Dhingra and Morrow (2019) in order to highlight the novel role played by the presence of off-market time in preferences. There is a continuum of potential firms that can each produce a unique variety of consumption good. Firms are ex-ante identical and must pay a fixed cost f_e in order to draw a marginal cost κ from some distribution G,

---

[14]By the mean-value theorem and the concavity of ψ, for any $p > 0$ there exists $\underline{p} \in (0, p)$ such that $\psi(p) = \psi(0) + p\psi'(\underline{p}) \geq 0 + p\psi'(p)$, which gives $\epsilon_\psi(p) \leq 1$.

[15]This example may be interpreted as one in which the "production technology" is approximately Leontief, assuming that each play requires (roughly) the same amount of time to attend.

[16]In the above theatre example this second effect is likely negligible, because utility is presumably not increased by staying in the theatre beyond the completion of the play.

where G has a continuous, positive density, denoted g, defined on the set $\mathcal{K}$. Upon drawing their marginal cost, the firms must pay an additional fixed cost f in order to produce. Both of these fixed costs are denominated in units of effective labor and are interpreted as real resource costs associated with setting up a firm.

Entry costs and technology. The mass of firms that pay f_e and draw κ is denoted by M_e. Some fraction of these firms remain in the market and produce, making non-negative profits. We denote this latter mass of firms by $M \leq M_e$. Further, let $\mathcal{K} = [\underline{\kappa}, \overline{\kappa}] \subseteq \mathcal{I}$ denote the set of operating firms. We index firms by their marginal cost κ and refer to this as the firm's "type." The output of a firm of type κ that employs ℓ units of effective labor is $y(\kappa, \ell) = \ell/\kappa$, so that κ is the marginal cost of the firm and $1/\kappa$ is the firm's labor productivity. The labor supply, ℓ, may vary across firms. In both the efficient and equilibrium allocations defined below, the set of marginal costs of firms that produce is of the form $[\underline{\kappa}, \overline{\kappa}]$ for some $\overline{\kappa} > \underline{\kappa}$. The cutoff value $\overline{\kappa}$ is an equilibrium object determined by the profit-maximizing decisions of firms. We refer to the determination of this value of $\overline{\kappa}$ as firm "selection."

Assumption 2.5 (Indexing). *The set indexing consumption experiences, $\mathcal{I}$, is identical to the set indexing producers' marginal costs, $\mathcal{K}$, so each q_i is associated with exactly one firm output, $q(\kappa) := y(\kappa, \ell)$, and each n_i can be written $n(\kappa)$.*

Assumption 2.5 is a natural assumption given that we interpret n_i as the time devoted to the consumption of the ith variety, and the variety is assumed to not exist until the associated firm pays the necessary fixed costs in order to enter and operate.[17] We therefore define the market equilibrium and efficient allocations below under Assumption 2.5.[18] We can therefore index market consumption and time use by the same variable $\kappa \in [\underline{\kappa}, \overline{\kappa}]$ which indexes the firms.

An entering firm has paid fixed cost f_e to draw a marginal cost κ. It must then pay an addition fixed cost f in order to operate (i.e. produce positive output). The problem of a firm who has paid both fixed costs is then defined as follows.

Definition 2.6 (Firm problem and operating profits). Given a multiplier λ on the consumer's budget constraint and the associated demand schedule $q(p; \lambda)$, an operating firm of type κ chooses p to solve

$$\pi(\kappa) = \max_{p \geq 0} (p - \kappa) q(p; \lambda) - f. \tag{2.7}$$

The quantity $\pi(\kappa)$ in (2.7) will be referred to as the operating profits of the firm.

---

[17]This assumption is consistent with the usual approach adopted in models of endogenous firm variety without time use (see, e.g., Dhingra and Morrow (2019)).

[18]In Section 4.1 we discuss how Assumption 2.5 may be relaxed to allow for consumers to derive utility from off-market activities not associated with the existence of a good produced by a firm.

Markups and holistic markups. The markup ratio of a firm of type κ that chooses price p is defined as $m(\kappa, p) := p/\kappa$, and the choice of the firm $p(\kappa)$ solves

$$m(\kappa, p(\kappa)) = 1 - \frac{1}{\epsilon_q(p(\kappa)) + 1}. \tag{2.8}$$

The firm-level markup for a variety is the ratio of the market price to its marginal cost of production. However, market production is not the only resource cost incurred in the production of final consumption, because consumers also devote time to off-market activities. This motivates an alternative markup concept, which we call "holistic markups," that represents the price of final consumption paid by the consumer divided by its marginal resource cost. To this end, for any $\kappa, p \geq \underline{\kappa}$ we define

$$\phi(\kappa, p) := \frac{\kappa + \Gamma(p)}{h(1, \Gamma(p))} \tag{2.9}$$

which represents the resources expended per unit of final consumption of a variety κ when the price is p, and define the holistic markup ratio $\mu(\kappa, p)$ by

$$\mu(\kappa, p) := \frac{\psi(p)}{\phi(\kappa, p)}. \tag{2.10}$$

An application of the envelope theorem to (2.5) then gives

$$\phi(\kappa, p) = \psi(p) - \psi'(p)(p - \kappa) \tag{2.11}$$

which gives the following relationship between the above markup notions.

Lemma 2.7. *For all $p, \kappa > 0$, the holistic markup ratio satisfies*

$$\mu(\kappa, p) = \frac{m(\kappa, p)}{\epsilon_\psi(p) + (1 - \epsilon_\psi(p))m(\kappa, p)}. \tag{2.12}$$

Further, when $p \geq \kappa$ we have $1 \leq \mu(\kappa, p) \leq m(\kappa, p)$, with $1 = \mu(\kappa, p) = m(\kappa, p)$ if and only if $\epsilon_\psi(p) = 1$ or $p = \kappa$.

Proof. The expression (2.12) follows from the definition (2.10) and equation (2.11), while the inequality $\mu(\kappa, p) \leq m(\kappa, p)$ follows from the concavity of ψ, because this implies that $\epsilon_\psi(p) \leq 1$. □

Lemma 2.7 shows that the firm-level markup $m(\kappa, p)$ always (weakly) overstates the above holistic equilibrium markup on final consumption, regardless of the preferences over final consumption.

2.3 Equilibrium and efficient allocations

We now turn to formal definitions of equilibrium and efficient allocations and the relationship between the two. We denote by M_e the mass of firms that enter and draw a marginal cost, so that the total effective labor used in production is $\mathcal{L} = M_e \int_{\mathcal{K}} \ell(\kappa) G(\mathrm{d}\kappa)$ and the total entry costs paid by firms are $\mathcal{E} = (f_e + fG(\overline{\kappa})) M_e$. Because consumers are identical and have unit mass and labor is the only input used in production, the individual time-use constraint (2.3) is also the aggregate resource constraint. Under Assumption 2.5 total off-market time is given by $\mathcal{N} = M_e \int_{\mathcal{K}} n(\kappa) G(\mathrm{d}\kappa)$ and the aggregate resource constraint may be written

$$M_e \left[\int_{\underline{\kappa}}^{\overline{\kappa}} (\kappa q(\kappa) + n(\kappa) + f) G(\mathrm{d}\kappa) + f_e \right] = \overline{T}. \tag{2.13}$$

Given the environment described in Sections 2.1 and 2.2, the following definitions of equilibrium and efficient allocations are standard extensions of their analogues with no off-market time.

Definition 2.8 (Equilibrium). A monopolistically competitive equilibrium consists of a cutoff $\overline{\kappa}^*$, a mass of firms M_e^*, market quantities $\mathbf{q}^* = \{q^*(\kappa)\}_{\kappa \in [\underline{\kappa}, \overline{\kappa}^*]}$, off-market time $\mathbf{n}^* = \{n^*(\kappa)\}_{\kappa \in [\underline{\kappa}, \overline{\kappa}^*]}$, and prices $\mathbf{p}^* = \{p^*(\kappa)\}_{\kappa \in [\underline{\kappa}, \overline{\kappa}^*]}$, such that:

(i) the cutoff $\overline{\kappa}^*$ satisfies $\pi(\overline{\kappa}^*) = 0$;

(ii) for all $\kappa \in [\underline{\kappa}, \overline{\kappa}^*]$, a firm of type κ chooses $p^*(\kappa)$ to solve (2.7);

(iii) given prices $\mathbf{p}^*$, $\mathbf{q}^*$ and $\mathbf{n}^*$ solve the consumer problem (2.4);

(iv) given $\overline{\kappa}^*$ and $\mathbf{q}^*$, the mass of entering firms, M_e^*, ensures that market resources add up to total time on the market, $\mathcal{L}^* + \mathcal{E}^* = M_e^* \left[\int_{\underline{\kappa}}^{\overline{\kappa}^*} (\kappa q^*(\kappa) + f) G(\mathrm{d}\kappa) + f_e \right]$.

Definition 2.9 (Efficient allocation). An efficient allocation consists of a cutoff $\widetilde{\kappa}$, a mass of firms $\widetilde{M}_e$, market quantities $\widetilde{\mathbf{q}} = \{\widetilde{q}(\kappa)\}_{\kappa \in [\underline{\kappa}, \widetilde{\kappa}]}$, and off-market time $\widetilde{\mathbf{n}} = \{\widetilde{n}(\kappa)\}_{\kappa \in [\underline{\kappa}, \widetilde{\kappa}]}$, that maximizes (2.1) subject to the resource constraint (2.13).

Definition 2.9 defines an efficient allocation in terms of market consumption $\mathbf{q}$ and off-market time $\mathbf{n}$. Using the definition of ψ in (2.5), if $\widetilde{\kappa}$, $\widetilde{M}_e$ and $\widetilde{\mathbf{c}}$ solve the problem

$$\max_{\mathbf{c}, M_e, \overline{\kappa}} M_e \int_{\underline{\kappa}}^{\overline{\kappa}} u(c(\kappa)) G(\mathrm{d}\kappa)$$

$$M_e \left[\int_{\underline{\kappa}}^{\overline{\kappa}} (\psi(\kappa) c(\kappa) + f) G(\mathrm{d}\kappa) + f_e \right] = \overline{T}. \tag{2.14}$$

then we obtain an efficient allocation with cutoff $\widetilde{\overline{\kappa}}$ and mass of firms $\widetilde{M_e}$ by defining $\widetilde{q}(\kappa) = \psi'(\kappa)\widetilde{c}(\kappa)$ and $\widetilde{n}(\kappa) = (\psi(\kappa) - \psi'(\kappa)\kappa)\widetilde{c}(\kappa)$ for all $\kappa \in [\underline{\kappa}, \widetilde{\overline{\kappa}}]$. For brevity, we will sometimes refer to the solution to (2.14) as the efficient allocation, with the understanding that $\widetilde{\mathbf{q}}$ and $\widetilde{\mathbf{n}}$ are related to $\widetilde{\mathbf{c}}$ in this manner.

In view of Definition 2.8 and Definition 2.9, the relationship between equilibrium and efficient allocations is not obvious: the former is defined in terms of prices and quantities that solve optimization problems for firms and consumers and satisfy market-clearing conditions, while the latter is the solution to a single optimization problem. Before turning to our results on markups, selection and the aggregate allocation of resources, we first discuss how we intend to compare these two allocations.

To begin, we recall that in a model with inelastic time use, Dhingra and Morrow (2019) show that equilibrium allocations maximize the integral of $qu'(q)$ subject to the aggregate resource constraint, and are therefore efficient for all distributions G if and only if $\epsilon_u(q) = qu'(q)/u(q)$ is constant in q. In our environment, there is an additional reason why equilibrium and efficient allocations can differ, namely, that resource costs are higher in the equilibrium allocation because the consumer's time use depends endogenously on market prices. Proposition 2.10 shows this by proving that equilibrium allocations can be recovered from the solution to an optimization problem. To motivate the form of this problem, note that when λ^* is the multiplier on the consumer's budget constraint, prices must satisfy $p^*(\kappa) = \psi^{-1}(u'(c(\kappa))/\lambda^*)$. The following resource constraint is then stated in terms of the holistic marginal resource cost, $\phi(\kappa, p)$, evaluated at these prices.

Proposition 2.10. *If for some $\lambda^* > 0$, the multiplier in the optimization problem*

$$V(\lambda^*) = \max_{\mathbf{c}, M_e, \overline{\kappa}} M_e \int_{\underline{\kappa}}^{\overline{\kappa}} c(\kappa)u'(c(\kappa))G(\mathrm{d}\kappa)$$

$$M_e\left[\int_{\underline{\kappa}}^{\overline{\kappa}} \left(\phi\big(\kappa, \psi^{-1}(u'(c(\kappa))/\lambda^*)\big)c(\kappa) + f\right)G(\mathrm{d}\kappa) + f_e\right] = \overline{T} \tag{2.15}$$

equals λ^*, then the associated allocation constitutes an equilibrium allocation in which for all $\kappa \in [\underline{\kappa}, \overline{\kappa}^*]$, prices are $p^*(\kappa) = \psi^{-1}(u'(c^*(\kappa))/\lambda^*)$, market consumption is $q^*(\kappa) = \psi'(p^*(\kappa))c^*(\kappa)$ and off-market time is $n^*(\kappa) = (\psi(p^*(\kappa)) - \psi'(p^*(\kappa))p^*(\kappa))c^*(\kappa)$.

Proof. See Appendix B.1. $\square$

Proposition 2.10 implies that for each $\kappa \in [\underline{\kappa}, \overline{\kappa}^*]$, the equilibrium $c^*(\kappa)$ maximizes equilibrium surplus, defined as

$$S^*(\kappa, \lambda^*) := \sup_{c>0} cu'(c) - \lambda^*\phi\big(\kappa, \psi^{-1}(u'(c)/\lambda^*)\big)c, \tag{2.16}$$

14

while if $\widetilde{\lambda}$ is the multiplier on the resource constraint in (2.14), then for each $\kappa \in [\underline{\kappa}, \widetilde{\overline{\kappa}}]$, the efficient $\widetilde{c}(\kappa)$ maximizes social surplus, defined as

$$\widetilde{S}(\kappa, \widetilde{\lambda}) := \sup_{c>0} u(c) - \widetilde{\lambda}\psi(\kappa)c. \tag{2.17}$$

The key point in the above is that in both the equilibrium and efficient allocation, the cutoff $\overline{\kappa}$ and mass M_e solve a problem of the form

$$\max_{M_e, \overline{\kappa}} M_e \left[\int_{\underline{\kappa}}^{\overline{\kappa}} S(\kappa, \lambda)G(\mathrm{d}\kappa) - \lambda(fG(\overline{\kappa}) + f_e) \right] \tag{2.18}$$

for some $\lambda > 0$ and function S. For the equilibrium allocation, S is defined in (2.16), while in the efficient allocation, S is defined in (2.17). The representation (2.18) shows that both the efficient and equilibrium allocations may be viewed as the outcome of choosing final consumption c, a cutoff value $\overline{\kappa}$, and mass of firms M_e to maximize some notion of surplus minus entry costs per variety and in general differ for two conceptually distinct reasons. First, as emphasized in Dhingra and Morrow (2019), the first term in the efficient surplus is utility, $u(c)$, while the first term in the equilibrium surplus is instead $cu'(c)$, which is proportional to utility only in the special case of CES preferences. Second, and novel to this paper, the resource cost per unit of consumption in the efficient allocation is $\psi(\kappa)$, while the corresponding resource cost in the equilibrium allocation is instead the (endogenous) quantity $\phi(\kappa, \psi^{-1}(u'(c)/\lambda^*)) \geq \psi(\kappa)$.

2.4 Firm selection and aggregate profits

The fact that firm entry is a two-step process implies that there are two margins along which we can compare equilibrium and efficient allocations: the productivity of the least productive firm that operates and the mass of firms that draw a marginal cost. We refer to the first comparison as the study of firm "selection," and we address the second comparison by characterizing the aggregate resources devoted to firm entry, $M_e(f_e + fG(\overline{\kappa}))$ (which equals operating profits in equilibrium), in both the efficient and the equilibrium allocations.

Selection. By combining the first-order conditions with respect to $\overline{\kappa}$ and M_e in the problem (2.18), we see that when an interior solution for $\overline{\kappa}$ exists, it solves

$$\frac{S(\overline{\kappa}, \lambda)}{\int_{\underline{\kappa}}^{\overline{\kappa}} S(\kappa, \lambda)G(\mathrm{d}\kappa)} = \frac{f}{fG(\overline{\kappa}) + f_e}. \tag{2.19}$$

In view of equation (2.19), the marginal cost $\overline{\kappa}$ of the least productive operating firm is

determined by equating the elasticity of the average surplus per firm, $\int_{\underline{\kappa}}^{\overline{\kappa}} S(\kappa, \lambda) G(\mathrm{d}\kappa)$, with the elasticity of the average entry cost per firm, $fG(\overline{\kappa}) + f_e$.

Aggregate resources. If we define $C^*(\kappa, \lambda^*) := \lambda^* \phi(\kappa, \psi^{-1}(u'(c^*(\kappa))/\lambda^*)) c^*(\kappa)$ and $\widetilde{C}(\kappa, \widetilde{\lambda}) := \widetilde{\lambda} \psi(\kappa) \widetilde{c}(\kappa)$, then the aggregate resources devoted to firm entry can be obtained by combining the first-order condition for M_e in the problem (2.18) with the resource constraint to obtain

$$M_e(fG(\overline{\kappa}) + f_e) = \frac{\overline{T} \int_{\underline{\kappa}}^{\overline{\kappa}} S(\kappa, \lambda) G(\mathrm{d}\kappa)}{\int_{\underline{\kappa}}^{\overline{\kappa}} (1 - 1/\nu(\kappa, \lambda))^{-1} S(\kappa, \lambda) G(\mathrm{d}\kappa)}. \tag{2.20}$$

where we wrote

$$\nu(\kappa, \lambda) = \frac{S(\kappa, \lambda)}{C(\kappa, \lambda)} + 1, \tag{2.21}$$

which represents the extent to which the surplus associated with variety κ exceeds its resource cost, and so may be interpreted as a kind of markup. Indeed, note that by equation (2.16), we have

$$\nu^*(\kappa, \lambda^*) = \mu(\kappa, p^*(\kappa)) \tag{2.22}$$

where $p^*(\kappa) = \psi^{-1}(u'(c^*(\kappa))/\lambda^*)$, and so $\nu^*(\kappa, \lambda^*)$ is the holistic markup (2.10) evaluated at equilibrium prices.

Equations (2.19) and (2.20) govern the cutoff value $\overline{\kappa}$ and the mass of firms M_e in both the equilibrium and efficient allocations. The key point of this section is that the form of these equations is the same across the efficient and equilibrium allocations, and so these allocations differ only insofar as the functions S and ν differ.

3 Parameterized model

In this section we will impose functional forms on the utility function u and the home production function h in order to sharpen these characterizations.

Assumption 3.1 (CES utility). *For some $\rho \in (0, 1)$ we have $u(c) = c^\rho$ for all $c \geq 0$.*

Assumption 3.2 (CES production). *For some $\alpha \in (0, 1)$ and $\xi \in [-\infty, 1]$ we have $h(q, n) = (\alpha q^\xi + (1 - \alpha) n^\xi)^{1/\xi}$ for all $q, n \geq 0$, where for $\xi = -\infty$ we define $h(q, n) = \min\{q, n\}$ and for $\xi = 0$ we define $h(q, n) = q^\alpha n^{1-\alpha}$.*

In Assumption 3.2, α is the weight of market consumption in final consumption and $(1 - \xi)^{-1}$ is the elasticity of substitution between market consumption and off-market time. We refer to $\xi = -\infty$ as the perfect complements (or Leontief) case,

and refer to $\xi = 0$ and $\xi = 1$ as the Cobb-Douglas and perfect substitutes cases, respectively. For $\xi \in (-\infty, 0) \bigcup (0, 1)$, the holistic price function becomes

$$\psi(p) = \frac{p}{\alpha}\left(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{1}{1/\xi - 1}}\right)^{1-1/\xi} \tag{3.1}$$

while $\psi(p) = p + 1$ for $\xi = -\infty$, $\psi(p) = \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}p^\alpha$ for $\xi = 0$, and $\psi(p) = \min\{p/\alpha, 1/(1-\alpha)\}$ for $\xi = 1$. Further, we have the following explicit expression for the associated elasticity.

Lemma 3.3 (Production elasticity). *For $\xi \notin \{-\infty, 0, 1\}$, the elasticity of ψ is*

$$\epsilon_\psi(p) = \frac{1}{1 + (1/\alpha - 1)^{\frac{1}{1-\xi}} p^{\frac{\xi}{1-\xi}}}.$$

For $\xi = -\infty$, $\epsilon_\psi(p) = p/(p+1)$, for $\xi = 0$, $\epsilon_\psi(p) = \alpha$, and for $\xi = 1$, $\epsilon_\psi(p) = 1$ for $p(1/\alpha - 1) \leq 1$ and is zero otherwise. Further, $\epsilon_{\psi'}(p) = (\epsilon_\psi(p) - 1)/(1 - \xi)$.

Proof. See Appendix B.2. $\qquad \square$

Lemma 3.3 shows that while the monotonicity of the holistic price of consumption is independent of ξ, the monotonicity of the elasticity is not. Indeed, whether the elasticity of the holistic price increases or decreases in the market price, p, depends solely upon whether off-market time and market goods are complements or substitutes. For the above specifications of preferences and technology, the price elasticities of market consumption and time use then simplify as follows.

Lemma 3.4. *Under Assumptions 3.1 and 3.2, the price elasticity of time use is*

$$\epsilon_n(p) = \left(\frac{1}{1-\xi} - \frac{1}{1-\rho}\right)\epsilon_\psi(p) \tag{3.2}$$

and the price elasticity of market consumption is

$$\epsilon_q(p) = -\frac{1}{1-\rho} + \left(\frac{1}{1-\xi} - \frac{1}{1-\rho}\right)(\epsilon_\psi(p) - 1). \tag{3.3}$$

For $p > 0$, we have $\epsilon_q(p) \geq -(1-\rho)^{-1}$ if $\xi \leq \rho$ and $\epsilon_q(p) \leq -(1-\rho)^{-1}$ if $\xi \geq \rho$.

Proof. See Appendix B.2. $\qquad \square$

For general u and ψ, Lemma 2.4 shows how the price elasticity of q can be decomposed into the price elasticities of c and q/c, respectively. Lemma 3.4 sharpens this by showing that for the above functional forms, this expression simplifies and

17

becomes affine in the elasticity of ψ. To proceed further in the characterization of markups, note that for $\xi \notin \{-\infty, 0, 1\}$ we can write the first-order condition of the firm as $F(p; \kappa) = 0$, where

$$F(p; \kappa) = \frac{1}{p/\kappa - 1} - \frac{1}{1/\rho - 1} + \left(\frac{1}{1/\rho - 1} - \frac{1}{1/\xi - 1} \right)(1 - \epsilon_\psi(p)). \qquad (3.4)$$

Because the function F is not necessarily monotonic in p, the existence or uniqueness of a solution to the firm's first-order condition is not obvious. Proposition 3.5 below shows that that a unique solution always exists and derives comparative statics for different values of ξ. When combined with Lemma 3.3, the expression on the right-hand side of equation (3.4) shows that there are three regions for ξ that govern the comparative statics for markups, because the coefficient of $\epsilon_\psi(p)$ in equation (3.4) is positive if and only if $\xi > \rho$, while $\epsilon_\psi(p)$ is increasing if and only if $\xi > 0$. This leads to closed-form expressions for markups when $\xi \in \{-\infty, 0, \rho\}$, and simple comparative statics for all other parameters.

Proposition 3.5. *Under Assumptions 3.1 and 3.2, there exists a unique solution to the firm's first-order condition, and the associated markup ratio $p(\kappa)/\kappa$ is constant in κ when $\xi \in \{0, \rho\}$, decreasing in κ when $\xi \in [-\infty, 0) \bigcup (\rho, 1)$, and increasing in κ when $\xi \in (0, \rho)$. Further, for $\xi \in \{-\infty, 0, \rho\}$:*

(i) (Leontief) For $\xi = -\infty$ we have $p(\kappa)/\kappa - 1 = (1/\rho - 1)(1 + 1/\kappa)$.

(ii) (Cobb-Douglas) For $\xi = 0$, we have $p(\kappa)/\kappa - 1 = (1/\rho - 1)/\alpha$.

(iii) For $\xi = \rho$, we have $p(\kappa)/\kappa - 1 = 1/\rho - 1$.

The markup ratio weakly exceeds $1/\rho$ if $\xi \le \rho$, is bounded above by $1/\rho$ when $\xi \ge \rho$, and equals $1/\rho$ when $\xi = \rho$.

Proof. See Appendix B.2. $\qquad \qquad \Box$

We emphasize that Proposition 3.5 characterizes the familiar markup ratio: the price charged by the firm divided by their marginal cost of production. In general, this ratio will differ from both of the markup notions defined in equation (2.21), which pertained to final consumption, and incorporated the value of time in the definition of marginal cost. By Lemma 2.7, we see that in general, the holistic equilibrium markup will be lower than the markup on market consumption. For our parameterized model, we may obtain sharper results.

Under Assumption 3.1, we can change variables to $p^* = \psi^{-1}(u'(c^*)/\lambda^*)$ and write the equilibrium surplus as

$$S^*(\kappa, \lambda^*) = \rho^{\frac{1}{1-\rho}}(\lambda^*)^{-\frac{\rho}{1-\rho}} \sup_{p^*>0} (p^* - \kappa)\psi'(p^*)\psi(p^*)^{-\frac{1}{1-\rho}} \tag{3.5}$$

while for social surplus, we change variables to $\widetilde{p} = \psi^{-1}(u(\widetilde{c})/[\widetilde{\lambda}\widetilde{c}])$ to write

$$\widetilde{S}(\kappa, \widetilde{\lambda}) = \widetilde{\lambda}^{-\frac{\rho}{1-\rho}} \sup_{\widetilde{p}>0} (\psi(\widetilde{p}) - \psi(\kappa))\psi(\widetilde{p})^{-\frac{1}{1-\rho}}. \tag{3.6}$$

Using the representations (3.5) and (3.6) of the equilibrium and social surplus, we may show the following.

Lemma 3.6. *Under Assumptions 3.1 and 3.2, for all $\kappa \geq \underline{\kappa}$ the equilibrium and efficient quantities in (2.21) are independent of the multipliers and satisfy $\nu^*(\kappa) \leq 1/\rho = \widetilde{\nu}(\kappa)$, with equality only in the case of perfect complements ($\xi = -\infty$).*

Proof. The first-order condition for (3.5) implies

$$0 = \frac{p^*}{p^* - \kappa} + \epsilon_{\psi'}(p^*) - \frac{\epsilon_\psi(p^*)}{1 - \rho} \tag{3.7}$$

while $\nu^*(\kappa) \leq 1/\rho$ is equivalent to $p^*/(p^* - \kappa) - \epsilon_\psi(p^*)/(1 - \rho) \geq 0$, which follows from equation (3.7) and the fact that $\epsilon_{\psi'} \leq 0$ with equality only when $\xi = -\infty$. The identity $\widetilde{\nu}(\kappa) = 1/\rho$ then follows directly from the first-order condition for (3.6). $\square$

When combined with equation (2.22), Lemma 3.6 shows that equilibrium holistic markups coincide with the efficient holistic markups only when $\xi = -\infty$. For all interior elasticities of substitution for q and n, efficient holistic markups weakly exceed equilibrium holistic markups, despite being weakly less than the equilibrium markups by Lemma 2.7.

3.1 Selection and aggregate entry

We now turn to the analysis of firm selection for our parameterized model. The key parameter governing our results is again the elasticity of substitution between off-market time and demand for market varieties, which is governed by ξ. As shown in Proposition 3.5, this elasticity dictates how markups vary in firm costs. We now show that this parameter also governs the efficiency or inefficiency of firm selection.

To this end, we first recall from equation (2.19) that in both the equilibrium and efficient allocations, the marginal cost $\overline{\kappa}$ of the least productive operating firm

is determined by equating the elasticity of the average surplus per firm with the elasticity of the average entry cost per firm. The proof of Proposition 3.7 first proceeds from this observation to show that characterization equilibrium and efficient selection reduces to comparing the elasticities of the two surplus functions in (3.5) and (3.6). Together with an application of the envelope theorem, this in turn reduces to signing the magnitude of equilibrium markups and gives the following.

Proposition 3.7 (Selection). *The equilibrium $\overline{\kappa}$ is inefficiently high if $\xi \in (-\infty, 0)$, inefficiently low if $\xi \in (0, 1)$, and efficient if $\xi \in \{-\infty, 0\}$.*

Proof. First note that if $\hat{S}$ and $\overline{S}$ are two smooth, positive functions on $[\underline{\kappa}, \infty)$ satisfying $\epsilon_{\hat{S}}(\kappa) \geq \epsilon_{\overline{S}}(\kappa)$ for all $\kappa \geq \underline{\kappa}$, then $\hat{S}(\kappa)/\overline{S}(\kappa)$ is weakly increasing and so

$$\int_{\underline{\kappa}}^{\overline{\kappa}} \hat{S}(\kappa)G(\mathrm{d}\kappa) \leq \int_{\underline{\kappa}}^{\overline{\kappa}} \hat{S}(\kappa)\left(\frac{\overline{S}(\kappa)}{\hat{S}(\kappa)}\frac{\hat{S}(\overline{\kappa})}{\overline{S}(\overline{\kappa})}\right)G(\mathrm{d}\kappa) = \frac{\hat{S}(\overline{\kappa})}{\overline{S}(\overline{\kappa})}\int_{\underline{\kappa}}^{\overline{\kappa}} \overline{S}(\kappa)G(\mathrm{d}\kappa) \qquad (3.8)$$

for all $\overline{\kappa} \geq \underline{\kappa}$. In view of equation (2.19) and the inequality (3.8), it will suffice to show that for all $\kappa \geq \underline{\kappa}$, we have $\epsilon_{S^*}(\kappa) \geq \epsilon_{\widetilde{S}}(\kappa)$ if $\xi \leq 0$, $\epsilon_{S^*}(\kappa) \leq \epsilon_{\widetilde{S}}(\kappa)$ if $\xi \geq 0$, and equality if $\xi \in \{-\infty, 0\}$. Applying the envelope theorem to expressions (3.5) and (3.6) gives $\epsilon_{S^*}(\kappa) = -(p^*(\kappa)/\kappa - 1)^{-1}$ and $\epsilon_{\widetilde{S}}(\kappa) = -\epsilon_\psi(\kappa)(1/\rho - 1)^{-1}$, respectively, and so if $\hat{p}(\kappa) := \kappa(1 + (1/\rho - 1)/\epsilon_\psi(\kappa))$ then $\epsilon_{S^*}(\kappa) \geq \epsilon_{\widetilde{S}}(\kappa)$ is equivalent to

$$F(\hat{p}(\kappa), \kappa) = \frac{\epsilon_\psi(\kappa) - 1}{1/\rho - 1} + \left(\frac{1}{1/\rho - 1} - \frac{1}{1/\xi - 1}\right)(1 - \epsilon_\psi(\hat{p}(\kappa))) \geq 0.$$

This in turn is equivalent to $H(\kappa) \leq 0$, where

$$H(\kappa) := \frac{1 - \epsilon_\psi(\kappa)}{1 - \epsilon_\psi(\hat{p}(\kappa))} + \frac{1/\rho - 1}{1/\xi - 1} - 1.$$

Using Lemma 3.3, the function $j(p) := (1 - \epsilon_\psi(p))^{-1} - 1$ is concave when $\xi \leq 0$, and so by considering a linear approximation of $j(\hat{p}(\kappa))$ about κ and simplifying, we have

$$\begin{aligned}
H(\kappa) &\leq \frac{(\hat{p}(\kappa) - \kappa)j'(\kappa)}{j(\kappa) + 1} + \frac{1/\rho - 1}{1/\xi - 1} \\
&= -\frac{(\hat{p}(\kappa)/\kappa - 1)}{1/\xi - 1}\epsilon_\psi(\kappa) + \frac{1/\rho - 1}{1/\xi - 1} = 0
\end{aligned} \qquad (3.9)$$

with the reverse inequality for $\xi \geq 0$, which gives the results for $\xi \neq -\infty, 0$. Finally, selection is efficient when $\xi = -\infty, 0$ because in these cases the function j is either constant or linear and the inequality (3.9) becomes an equality. $\qquad\square$

20

We believe that it is intuitive why equilibrium selection is efficient in the case of perfect complements. Because the ratio of off-market time to market consumption is independent of prices, this case is essentially a disguised version of a model with CES preferences and no off-market time, in which the marginal cost of production for a firm of type κ is now $\psi(\kappa) = \kappa + 1$. To see how this relates to the above algebra, note that in this case $\psi'(\kappa) = 1$ for all κ, and so the equilibrium and social surplus in (2.16) and (2.17) are proportional to one another. This is the sole case in which firms' market power does not distort the ratio n/q away from its efficient value.

Proposition 3.7 shows that as we increase the elasticity of substitution between time and consumption above zero, selection is first too lax (for $\xi < 0$) and then too strict (for $\xi > 0$). To understand why, note that when market goods and off-market time are complements, consumers cannot easily substitute away from market goods by reallocating time, and so equilibrium surplus (profits) declines more slowly than social surplus. High-cost firms can therefore profitably operate despite their inefficiency, as consumers still need to purchase market goods to utilize their time, which leads to (weakly) inefficiently high entry of high-cost firms. Conversely, when market goods and time are substitutes, consumers can more easily respond to high prices by substituting time for market purchases. This heightened elasticity of substitution ensures that as costs rise, profits fall more rapidly than social surplus, and prevents some marginally efficient high-cost firms from operating.

We now study aggregate resources devoted to entry in our parameterized model. Recall that the first entry cost, f_e, is paid by all of the entering firms, while the second entry cost, f, is paid only by the fraction $G(\overline{\kappa})$ that choose to operate. The total resources expended in the process of firm creation are then $(f_e + fG(\overline{\kappa}))M_e$. Proposition 3.8 below shows that the equilibrium value of this quantity is always (weakly) below its efficient counterpart, and coincides in one special case.

Proposition 3.8 (Inefficiently low entry)**.** *The total resources devoted to firm entry, $\mathcal{E} = M_e(f_e + fG(\overline{\kappa}))$, are inefficiently low in equilibrium unless off-market time and market goods are perfect complements ($\xi = -\infty$).*

Proof. This is immediate from equation (2.20) and Lemma 3.6. $\qquad\square$

Proposition 3.8 shows that outside of the special case of perfect complementarity between off-market time and market goods, the equilibrium allocation devotes too few resources to firm creation. Note also that due to the assumption of free entry (and hence zero ex-ante profits), in equilibrium the quantity $M_e(f_e + fG(\overline{\kappa}))$ coincides with the aggregate ex-post operating profits of all active firms. Proposition 3.8 shows that in our model structure these profits are typically "too small" to provide incentives for efficient firm entry. Intuitively, consumers in this economy do not internalize the fact

that their decisions to allocate labor between off-market time and market production affect firms' profits and the incentives to enter. This is an additional economic force that is not present in models without off-market time.

Proposition 3.7 characterizes when the efficient and equilibrium cutoff values of $\overline{\kappa}$ coincide, while Proposition 3.8 characterizes when the efficient and equilibrium values of $(f_e + fG(\overline{\kappa}))M_e$ coincide. Combining these two results, we obtain the following.

Corollary 3.9 (Efficiency). *The equilibrium allocation is efficient if and only if off-market time and market goods are perfect complements in the consumption technology.*

The efficiency of the equilibrium allocation under perfect complements is particularly noteworthy because Proposition 3.5 shows that the (usual) markups in this case are not constant across firms. An economist who were to observe this heterogeneity would therefore erroneously conclude that welfare losses existed in this environment if they ignored the role of off-market time. Note, however, that by Lemma 3.6, in this case the holistic markups are constant across firms. Empirical work seems to suggest an increasing relationship between markups and productivities.[19] The above example shows that such variable markups may, indeed, be efficient. Further, when combined with Proposition 3.5, Corollary 3.9 shows that when time use is elastically supplied and non-separable with market varieties, the link between markup constancy and efficiency is broken in both directions, in the sense that such constancy is neither necessary nor sufficient for efficiency. For instance, when $\xi = -\infty$, markups are variable across firms and the allocation is efficient, and so clearly constancy of markups is not necessary. On the other hand, when $\xi = \rho$, markups are constant across firms but selection is inefficiently low, and so such constancy is also not sufficient.

3.2 Firm-specific comparisons

Proposition 3.8 shows that outside of the case of perfect complements, the total amount of resources devoted to firm entry is inefficiently low in equilibrium. In view of the resource constraint (2.3), this means that the sum of the aggregate resources devoted to production and off-market time are inefficiently high.

It is then natural to ask whether the time use and market consumption associated with a particular variety are "too high" or "too low" from the point of view of the planner. In general, it is not possible to compare such quantities for all types of firms because the distributions of firm productivity differ across the efficient and equilibrium allocations. However, in the case of Cobb-Douglas technology, $\xi = 0$, Proposition 3.7 shows that the distribution of firms that operate in the equilibrium and efficient allocations coincide, and we can sharpen the above results.

---

[19]See, e.g., Edmond and Veldkamp (2009); Berry et al. (2019); Peters (2020).

Proposition 3.10. *When $u(c) \equiv c^\rho$ and $h(q, n) \equiv q^\alpha n^{1-\alpha}$ for some $\rho, \alpha \in (0, 1)$, the ratios of equilibrium consumption, time use and market consumption to their efficient counterparts are constant across firms and satisfy*

$$\frac{n^*}{\widetilde{n}} = 1 + (1/\rho - 1)/\alpha > \frac{\overline{c}^*}{\widetilde{\overline{c}}} = (1 + (1/\rho - 1)/\alpha)^{1-\alpha} > \frac{q^*}{\widetilde{q}} = 1 \qquad (3.10)$$

while $M_e^* < \widetilde{M}_e$.

Proof. See Appendix B.2. ☐

Relative to the efficient allocation, the equilibrium number of varieties is lower, the final consumption of each variety is higher, but the market production for each variety is unchanged. Consequently, firms are neither "too big" nor "too small," but there are too few of them in the aggregate, and consumers devote an inefficiently high amount of time to off-market activities. Thus, while market competition forces the right "kind" of firms to operate (i.e., selection is efficient), there are not enough competitors making market products, and consumers are not supplying enough labor toward creating new firms (i.e., $\mathcal{E}$).

4 Discussion

In this section we discuss the roles of various assumptions in our model and relate the results once again to the literature. Section 4.1 considers the effect of relaxing Assumption 2.5, while Section 4.2 discusses the role of non-separability and the extent to which the insights of this paper apply to other settings.

4.1 Allowing for leisure

Under Assumption 2.5 all time spent engaged in off-market activities requires the utilization of market varieties in order to generate final consumption experiences. In the modern world this assumption is not overly implausible: there are very few activities we engage in which do not also involve the utilization of a product bought and sold on the market. Even leisure time is usually complemented by the utilization of recreational goods and service: 1) walking in the park requires tennis shoes and perhaps a podcast streamed through Bluetooth headphones and paid for via a subscription service; 2) video gaming requires a console or computer, headsets and other gaming equipment, and the game software itself; 3) child care and personal care require a litany of consumer goods, like soaps, oils, wipes, etc.; 4) even socializing with one's family or friends is often done in conjunction with other consumption activities, like

eating out or eating at home, sipping coffee or wine, playing games, watching television or movies, etc. Indeed, it is difficult to come up with an example of an activity one engages in which is also not associated with the utilization of a product bought and sold on the market.

To show, however, that Assumption 2.5 is innocuous we relax it here, letting $\mathcal{K}$ be a strict subset of $\mathcal{I}$. We thus consider an economy where there exist $q(\kappa) > 0$ and associated consumption-time allocations $n(\kappa) > 0$ but also $n_i > 0$ where $i \notin \mathcal{K}$. Of course, $q_i = 0$ for $i \notin \mathcal{K}$. This requires us to restrict the set of possible functions h could take on to those which are finite when $q_i = 0$. Assumption 4.1 encodes this.

Assumption 4.1. *Let h be such that $h(0, n_i) > 0$ for $n_i > 0$, so that there exist final consumption experiences, c_i, for which consumers allocate some positive amount of their off-market time but for which no market goods/services are utilized.*

Under Assumption 4.1 and the form of preferences (2.1), we can write utility as the sum of utility from consumption experiences associated with using both market goods and time and consumption experiences associated only with using time, $\mathcal{U}(\mathbf{c}) = \int_{\mathcal{K}} u(c_i) \mathrm{d}i + \int_{\mathcal{I} \backslash \mathcal{K}} u(c_i) \mathrm{d}i$, and the budget constraint can be similarly split.

Segmenting the different kinds of experiences in this manner does not change the demand elasticities for market consumption or firms' pricing decisions. Indeed, the leisure bundle would be such that $\forall i, j \in \mathcal{I} \backslash \mathcal{K}$ such that $i \neq j$, $n_i = n_j$, and we could write $\int_{\mathcal{I} \backslash \mathcal{K}} u(c_i) \mathrm{d}i = \underline{N}$, where $\underline{N}$ is some leisure-utility index. The introduction of additively separated leisure utility thus does not affect the pricing decisions of firms.

4.2 Inferences from demand elasticities

In this paper we have characterized efficient and equilibrium allocations in an environment in which consumers exhibit love-for-variety and utility depends non-separably on both market consumption and off-market time. In this final section, we now relate the findings of this paper to those in the literature, in order to highlight the novel role of non-separable time use. Recall that Dhingra and Morrow (2019) consider an environment in which utility is independent of time use and exhibits variable elasticities of substitution (VES) across varieties. They show that the equilibrium $\bar{\kappa}$ is inefficiently high when ϵ_u is everywhere decreasing and inefficiently low when ϵ_u is everywhere increasing.[20] It is then natural to ask how inferences drawn from observed markups differ between our model and Dhingra and Morrow (2019).

To explore this in the simplest possible case, we first suppose that $h(q, n) \equiv q$ and $u(q) = q^\rho - \eta q$ for some $\eta > 0$, at least when restricted to $q < (\eta/\rho)^{-\frac{1}{1-\rho}}$. The

---

[20]See Proposition 5 on page 213 of Dhingra and Morrow (2019).

associated utility elasticity is then

$$\epsilon_u(q) = \rho - \frac{(1-\rho)\eta}{q^{\rho-1} - \eta} \qquad (4.1)$$

which is decreasing, and so by the results of Dhingra and Morrow (2019), the equilibrium $\overline{\kappa}$ is inefficiently high. However, the first-order condition rearranges to give the demand schedule $q(p) = (\lambda/\rho)^{-\frac{1}{1-\rho}}(p + \eta/\lambda)^{-\frac{1}{1-\rho}}$, which coincides with the demand schedules in the model of Section 3 with $h(q,n) = \min\{q,n\}$, $u(c) = c^\rho$ and $\lambda = \eta$. Consequently, the equilibrium allocation in the latter economy is efficient, while the equilibrium allocation in the former economy is not. In this example, an econometrician who ignored the role of time use and used markups to infer demand elasticities would therefore erroneously believe that there were welfare losses.

To see how this simple example generalizes, note that with inelastic time use, the demand schedule satisfies $u'(\hat{q}(p)) = \lambda p$, where λ is the multiplier on the consumer's budget constraint, and so using the change-of-variables $z = \hat{q}(w)$, we have

$$u(x) = \int_0^x u'(z)dz = \lambda \int_0^x \hat{q}^{-1}(z)dz = -\lambda \int_{\hat{q}^{-1}(x)}^\infty w\hat{q}'(w)dw \qquad (4.2)$$

for all $x \geq 0$. For this utility function, the associated elasticity satisfies

$$\epsilon_u(\hat{q}(p)) = \frac{p\hat{q}(p)}{\int_p^\infty (-w\hat{q}'(w))dw} \qquad (4.3)$$

and explicit calculation gives the following.

Lemma 4.2. *If $\epsilon_{\hat{q}}$ is decreasing everywhere, then ϵ_u is decreasing everywhere, where u is defined in equation (4.2).*

Proof. See Appendix B.2. □

Lemma 4.2 shows that if we interpreted the markups emerging from the model in Section 3 using the results of Dhingra and Morrow (2019), we would conclude that the equilibrium $\overline{\kappa}$ is inefficiently high when $\xi \in [-\infty, 0) \bigcup (\rho, 1)$, inefficiently low when $\xi \in (0, \rho)$ and efficient if $\xi \in \{0, \rho\}$. Note that this contrasts with our selection result Proposition 3.7, which shows that the equilibrium $\overline{\kappa}$ is inefficiently high if $\xi \in (-\infty, 0)$, inefficiently low if $\xi \in (0, 1)$, and efficient if $\xi \in \{-\infty, 0\}$. An econometrician who observed the markups generated by our model but ignored the role of off-market time would erroneously conclude that selection was not strict enough if $\xi \in (\rho, 1)$ (when in fact it is too strict) and erroneously conclude that it is not strict enough if $\xi = -\infty$ (when it is in fact efficient). Consequently, incorporating off-market time as we have

done in this paper does not merely change the distribution of markups but also the welfare implications of such markups.

Extensions to other environments. We have emphasized in this paper that the distribution of markups is not a sufficient statistic for inferring aggregate welfare when market consumption and off-market time enter non-separably into preferences. This non-separability ensures that the price paid to the firm does not represent the true economic price of final consumption. In this approach we have been motivated by the insights of Becker (1965), because this is an enrichment of the consumer problem that is likely familiar to most economists.

However, the above insight is likely applicable to other environments. For instance, suppose that consumers care solely about a single consumption good produced by a competitive sector of final goods producers, and that the technology of these producers is a CES function of a continuum of intermediate varieties. Suppose further that the production of each variety depends non-separably on both labor and the output of monopolistically competitive firms taking labor as their sole input. Then the situation would be similar to that analyzed in this paper, insofar as the markups of the intermediate producers would not represent the extent to which the price of a variety differed from its (total) marginal cost of production.

5 Conclusion

This paper has presented a model in which both the dispersion and welfare effects of markups depend upon the extent to which consumers value off-market time. In our model, the usual definition of markups differs from a more holistic definition that incorporates the fact that the price paid by consumers is not the sole economic cost they incur to consume a market good. For the special case of perfect complements between market goods and off-market time, we have shown that heterogeneous markups can be consistent with a first-best allocation.

We believe that our work ought to encourage researchers to continue to consider how accounting for different household decision structures affects inferences regarding efficiency when faced with market structures that allow for prices to exceed marginal costs. We have shown that there exists an economy in which variable markups are indeed efficient. The recent work of Parenti et al. (2017) represents a different enrichment of the standard consumer problem in which efficiency is also consistent with heterogeneous markups. It thus remains for researchers to determine which decision structures themselves are most plausible when assessing what such structures imply for the efficiency of environments in which firms operate with imperfect competition.

References

M. Aguiar and E. Hurst. Chapter 4 - The Macroeconomics of Time Allocation. In John B. Taylor and Harald Uhlig, editors, *Handbook of Macroeconomics*, volume 2, pages 203–253. Elsevier, January 2016. doi:10.1016/bs.hesmac.2016.06.004.

Mark Aguiar and Erik Hurst. Consumption versus Expenditure. *Journal of Political Economy*, 113(5):919–948, October 2005. doi:10.1086/491590.

Mark Aguiar and Erik Hurst. Life-Cycle Prices and Production. *American Economic Review*, 97(5):1533–1559, December 2007. doi:10.1257/aer.97.5.1533.

Mark Aguiar and Erik Hurst. Deconstructing Life Cycle Expenditure. *Journal of Political Economy*, 121(3):437–492, 2013. doi:10.1086/670740.

Mark Aguiar, Erik Hurst, and Loukas Karabarbounis. Recent Developments in the Economics of Time Use. *Annual Review of Economics*, 4(1):373–397, September 2012. doi:10.1146/annurev-economics-111809-125129.

Mark Aguiar, Mark Bils, Kerwin Kofi Charles, and Erik Hurst. Leisure luxuries and the labor supply of young men. *Journal of Political Economy*, 129(2):337–382, 2021.

Gary S. Becker. A Theory of the Allocation of Time. *The Economic Journal*, 75 (299):493–517, 1965. doi:10.2307/2228949.

W.L. Bednar and Nick Pretnar. Structural change with time to consume. *Working Paper*, 2024.

Kristian Behrens, Giordano Mion, Yasusada Murata, and Jens Suedekum. Quantifying the Gap Between Equilibrium and Optimum under Monopolistic Competition. *The Quarterly Journal of Economics*, 135(4):2299–2360, November 2020. doi:10.1093/qje/qjaa017.

Steven Berry, Martin Gaynor, and Fiona Scott Morton. Do Increasing Markups Matter? Lessons from Empirical Industrial Organization. *Journal of Economic Perspectives*, 33(3):44–68, August 2019. doi:10.1257/jep.33.3.44.

Florin O. Bilbiie, Fabio Ghironi, and Marc J. Melitz. Endogenous Entry, Product Variety, and Business Cycles. *Journal of Political Economy*, 120(2):304–345, April 2012. doi:10.1086/665825.

Florin O. Bilbiie, Fabio Ghironi, and Marc J. Melitz. Monopoly Power and Endogenous Product Variety: Distortions and Remedies. *American Economic Journal: Macroeconomics*, 11(4):140–174, October 2019. doi:10.1257/mac.20170303.

Corina Boar and Virgiliu Midrigan. Markups and Inequality. Working Paper 25952, National Bureau of Economic Research, June 2019.

Job Boerma and Loukas Karabarbounis. Inferring Inequality With Home Production. *Econometrica : journal of the Econometric Society*, 89(5):2517–2556, 2021. doi:10.3982/ECTA15966.

Benjamin Bridgman. Home productivity. *Journal of Economic Dynamics and Control*, 71:60–76, October 2016. doi:10.1016/j.jedc.2016.08.003.

Benjamin Bridgman, Georg Duernecker, and Berthold Herrendorf. Structural transformation, marketization, and household production around the world. *Journal of Development Economics*, 133:102–126, 2018a.

Benjamin Bridgman, Georg Duernecker, and Berthold Herrendorf. Structural transformation, marketization, and household production around the world. *Journal of Development Economics*, 133:102–126, July 2018b. doi:10.1016/j.jdeveco.2017.12.009.

Jan De Loecker, Jan Eeckhout, and Gabriel Unger. The Rise of Market Power and the Macroeconomic Implications*. *The Quarterly Journal of Economics*, 135(2):561–644, May 2020. doi:10.1093/qje/qjz041.

Jan De Loecker, Jan Eeckhout, and Simon Mongey. Quantifying market power and business dynamism in the macroeconomy. *NBER Working Paper #28761*, May 2021.

Swati Dhingra and John Morrow. Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity. *Journal of Political Economy*, 127(1):196–232, February 2019. doi:10.1086/700732.

Avinash K. Dixit and Joseph E. Stiglitz. Monopolistic Competition and Optimum Product Diversity. *The American Economic Review*, 67(3):297–308, 1977. ISSN 0002-8282.

Chris Edmond and Laura Veldkamp. Income dispersion and counter-cyclical markups. *Journal of Monetary Economics*, 56(6):791–804, September 2009. doi:10.1016/j.jmoneco.2009.06.006.

Chris Edmond, Virgiliu Midrigan, and Daniel Yi Xu. How Costly Are Markups? *Journal of Political Economy*, 131(7):1619–1675, July 2023. doi:10.1086/722986.

Paolo Epifani and Gino Gancia. Trade, markup heterogeneity and misallocations. *Journal of International Economics*, 83(1):1–13, January 2011. doi:10.1016/j.jinteco.2010.10.005.

Federico Etro. Research in economics and monopolistic competition. *Research in Economics*, 71(4):645–649, December 2017. doi:10.1016/j.rie.2017.10.009.

Lei Fang, Anne Hannusch, and Pedro Silos. Bundling time and goods: Implications for hours dispersion. *Working Paper*, 2022.

Zach Flynn, Amit Gandhi, and James Traina. Measuring markups with production data. Technical report, 2019.

Gilbert R. Ghez and Gary S. Becker. *The Allocation of Time and Goods over the Life Cycle*. Columbia University Press: New York, 1974.

Paul Gomme, Finn Kydland, and Peter Rupert. Home production meets time to build. *Journal of Political Economy*, 109(5), 2001.

Jeremy Greenwood and Zvi Hercowitz. The allocation of capital and time over the business cycle. *Journal of Political Economy*, 99(6):1188–1214, 1991.

Jeremy Greenwood, Ananth Seshadri, and Mehmet Yorukoglu. Engines of liberation. *Review of Economic Studies*, 72:109–133, 2005.

Reuben Gronau. Leisure, Home Production, and Work–the Theory of the Allocation of Time Revisited. *Journal of Political Economy*, 85(6):1099–1123, 1977. doi:10.1086/260629.

Robert E. Hall. New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy. Working Paper 24574, National Bureau of Economic Research, May 2018.

Rachel Ngai and Christopher Pissarides. Trends in hours and economic growth. *Review of Economic Dynamics*, 11:239–256, 2008.

Mathieu Parenti, Philip Ushchev, and Jacques-François Thisse. Toward a theory of monopolistic competition. *Journal of Economic Theory*, 167:86–115, January 2017. doi:10.1016/j.jet.2016.10.005.

Michael Peters. Heterogeneous Markups, Growth, and Endogenous Misallocation. *Econometrica : journal of the Econometric Society*, 88(5):2037–2073, 2020. doi:10.3982/ECTA15565.

Nick Pretnar. Measuring inequality with consumption time. *Working Paper*, 2024.

Evgeny Zhelobodko, Sergey Kokovin, Mathieu Parenti, and Jacques-François Thisse. Monopolistic Competition: Beyond the Constant Elasticity of Substitution. *Econometrica : journal of the Econometric Society*, 80(6):2765–2784, 2012. doi:10.3982/ecta9986.

A   Notation

In order to aid the reader, in this appendix we list the symbols used in this paper and provide a link to where they are first defined. We use star symbols $*$ for equilibrium quantities (e.g., the efficient market consumption of firm κ is $q^*(\kappa)$) and use large tildes for efficient allocations (e.g., the efficient market consumption of firm κ is $\widetilde{q}(\kappa)$).

- The utility function is denoted by u. The functional form for u adopted in the parameterized model is given in Assumption 3.1.

- The home production function is denoted by h. The functional form for h adopted in the parameterized model is given in Assumption 3.2.

- The holistic price function ψ is defined in equation (2.5).

- The operating profits of the firm are defined in Definition 2.6 and denoted π. These are net of the operating cost but not the entry cost.[21]

- The markup ratio (i.e. the usual notion of the markup) is defined in (2.8). We define this for κ (denoting the firm) and an arbitrary p (denoting the price). The equilibrium markup of firm κ is then $m^*(\kappa)$. In this environment, the firm can either be interpreted as choosing prices or quantities. We interpret it as choosing prices.

- Terminology regarding markups is not uniform in the literature. For both markup notions, We follow Hall (2018) and refer to p/κ as the "markup ratio," with the "markup" then defined as $p/\kappa - 1$.

---

[21]Our convention here follows Dhingra and Morrow (2019).

- The function $\phi(\kappa, p)$ represents the resources expended per unit of final consumption of a variety κ when the price is p and is defined in equation (2.9).

- The holistic markup ratio $\mu(\kappa, p)$ is defined in equation (2.10).

For any function f we write ϵ_f for the associated elasticity,

$$\epsilon_f(x) = \frac{x f'(x)}{f(x)}.$$

Note that in this paper we do not take absolute values for demand elasticities (as is sometimes done), and so such elasticities are negative.

B   Proofs

B.1   General results

Proof of Proposition 2.10. Suppose that for some multiplier $\lambda^* > 0$ the quantities $M_e^*, \overline{\kappa}^*$ and $\{c^*(\kappa)\}_{\kappa \in [\underline{\kappa}, \overline{\kappa}^*]}$ solve the problem (2.15) and that the associated multiplier on the constraint equals λ^*. We wish to use these quantities to construct an equilibrium allocation and prices. In so doing we will also show that λ^* is the multiplier on the consumer's budget constraint.

First, for any $q, \lambda > 0$, denote by $\hat{n}(q, \lambda)$ the time use chosen by the consumer given market consumption q and multiplier λ on the budget constraint, and note that

$$u'(h(q, \hat{n}(q, \lambda)))h_n(q, \hat{n}(q, \lambda)) = \lambda. \tag{B.1}$$

The definition of ϕ in equation (2.9) implies that for any $p, \kappa, q > 0$ we have

$$h(q, q\Gamma(p))\phi(\kappa, p) = \kappa q + q\Gamma(p),$$

and so substituting the price $p = \psi^{-1}(u'(h(q, \hat{n}(q, \lambda)))/\lambda)$, we have

$$h(q, \hat{n}(q, \lambda))\phi\big(\kappa, \psi^{-1}(u'(h(q, \hat{n}(q, \lambda)))/\lambda)\big) = \kappa q + \hat{n}(q, \lambda). \tag{B.2}$$

Now, for any $\kappa \in [\underline{\kappa}, \overline{\kappa}^*]$, by assumption, $c^*(\kappa)$ solves

$$S^*(\kappa, \lambda^*) := \sup_{c > 0} cu'(c) - \lambda^* \phi\big(\kappa, \psi^{-1}(u'(c)/\lambda^*)\big)c, \tag{B.3}$$

which means that the $q^*(\kappa)$ that solves $c^*(\kappa) = h(q^*(\kappa), \hat{n}(q^*(\kappa), \lambda^*))$ also solves

$$S^*(\kappa, \lambda^*) = \max_{q>0} h(q, \hat{n}(q, \lambda^*)) u'(h(q, \hat{n}(q, \lambda^*))) - \lambda^*(\kappa q + \hat{n}(q, \lambda^*)). \tag{B.4}$$

Because h exhibits constant returns to scale, we have $h(q, n) = q h_q(q, n) + n h_n(q, n)$ for all $q, n > 0$, and so (B.4) becomes

$$\begin{aligned} S^*(\kappa, \lambda^*) = \ &\max_{q>0} q h_q(q, \hat{n}(q, \lambda^*)) u'(h(q, \hat{n}(q, \lambda^*))) - \lambda^* \kappa q \\ &+ (h_n(q, \hat{n}(q, \lambda^*)) u'(h(q, \hat{n}(q, \lambda^*))) - \lambda^*) \hat{n}(q, \lambda^*) \end{aligned}$$

and the second term in parentheses vanishes by the definition of $\hat{n}(q, \lambda^*)$. Consequently, for any $\kappa \in [\underline{\kappa}, \overline{\kappa}^*]$, $q^*(\kappa)$ solves $\max_{q>0} p(q; \lambda^*) q - \kappa q$, where $p(\cdot; \lambda^*)$ is the inverse demand function faced by the firm given by

$$p(q; \lambda^*) = \frac{1}{\lambda^*} h_q(q, \hat{n}(q, \lambda^*)) u'(h(q, \hat{n}(q, \lambda^*))), \tag{B.5}$$

and so $q^*(\kappa)$ maximizes the operating profits π of the firm defined in Definition 2.6 and $S^*(\kappa, \lambda^*) = \lambda^*(\pi(\kappa) + f)$. It follows that $\overline{\kappa}^*$ and M_e^* solve

$$\lambda^* \max_{M_e, \overline{\kappa}} M_e \left[\int_{\underline{\kappa}}^{\overline{\kappa}} \pi(\kappa) G(\mathrm{d}\kappa) - f_e \right]. \tag{B.6}$$

The first-order condition in (B.6) with respect to $\overline{\kappa}$ ensures that the operating profit of the marginal firm is zero, and the first-order condition with respect to M_e ensures that firms obtain zero net profits in expectation, and so the above allocation and prices constitute an equilibrium. $\qquad \square$

B.2 Parameterized model

We now record the proofs for all claims pertaining to the parameterized model from Section 3. First recall the definition of h, $h(q, n) = (\alpha q^\xi + (1-\alpha) n^\xi)^{1/\xi}$ and note that the choice of n/q as a function of p is $\Gamma(p) = [(1/\alpha - 1)p]^{\frac{1}{1-\xi}}$. In this case, the holistic price function is

$$\psi(p; \alpha, \xi) = \frac{p}{\alpha} \left(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}} \right)^{1-1/\xi}. \tag{B.7}$$

Lemma B.1. *The derivative of the total cost function in* (B.7) *is given by*

$$\psi'(p; \alpha, \xi) = \left(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}} \right)^{-1/\xi}. \tag{B.8}$$

Proof. Explicit calculation gives

$$\psi'(p) = \frac{1}{\alpha}\Big(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}\Big)^{1-1/\xi}$$
$$- (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}\frac{1}{\alpha}\Big(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}\Big)^{-1/\xi}$$

which simplifies as claimed. $\qquad\square$

Proof of Lemma 3.3. To derive $\epsilon_\psi(p)$, we use Lemma B.1 to write

$$\frac{\psi(p)}{\psi'(p)} = p + (1/\alpha - 1)p[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}$$

from which the comparative statics regarding p and ξ follow. To derive $\epsilon_{\psi'}(p)$, we use Lemma B.1 to obtain

$$\psi''(p) = -\frac{1}{\xi}\Big(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}\Big)^{-1/\xi-1} \times \frac{\xi p^{-1}}{1-\xi}(1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}$$

and hence

$$\frac{p\psi''(p)}{\psi'(p)} = -\frac{1}{1-\xi}\Big(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}\Big)^{-1/\xi-1}$$
$$\times (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}\Big(\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}\Big)^{1/\xi} \qquad \text{(B.9)}$$
$$= -\frac{1}{1-\xi}\left(\frac{(1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}}{\alpha + (1-\alpha)[(1/\alpha - 1)p]^{\frac{\xi}{1-\xi}}}\right)$$

which simplifies as claimed. $\qquad\square$

Proof of Lemma 3.4. Using the functional form in Assumption 3.1, this follows immediately by combining Lemma 3.3 with the general expression (2.6). $\qquad\square$

Proof of Proposition 3.5. The proof proceeds in three steps: first, we establish existence of an optimal price; second, we prove uniqueness; and third, we derive comparative statics.

Step 1. Existence. Consider the firm's revenue as a function of price. This is bounded above by a multiple of $\epsilon_\psi(p)\psi(p)^{-\frac{\rho}{1-\rho}}$ and so profits vanish both when the firm equates its price with the marginal cost, $p = \kappa$, and as $p \to \infty$, because when $\xi \leq 0$, $\lim_{p\to\infty}\psi(p)^{-\frac{\rho}{1-\rho}} = 0$, and when $\xi > 0$, $\lim_{p\to\infty}\epsilon_\psi(p) = 0$. The optimal price

33

therefore satisfies the firm's first-order condition $F(p; \kappa) = 0$ for F given in

$$F(p; \kappa) = \frac{1}{p/\kappa - 1} - \frac{1}{1/\rho - 1} + \left(\frac{1}{1/\rho - 1} - \frac{1}{1/\xi - 1} \right)(1 - \epsilon_\psi(p)). \qquad \text{(B.10)}$$

Step 2. Uniqueness.

We establish uniqueness by showing $F'(p; \kappa) < 0$ whenever $F(p; \kappa) \geq 0$. Consider two cases:

Case 1: If $\xi \in (-\infty, 0) \bigcup (\rho, 1)$, then $F'(p; \kappa) < 0$ everywhere.

Case 2: For $\xi \in (0, 1)$ we first note that for any positive differentiable function f, the derivative of its elasticity satisfies:

$$p\epsilon'_f(p) = (1 + \epsilon_{f'}(p) - \epsilon_f(p))\epsilon_f(p) \qquad \text{(B.11)}$$

By Lemma 3.3, we know that:

$$\epsilon_{\psi'}(p) = \frac{\epsilon_\psi(p) - 1}{1 - \xi}. \qquad \text{(B.12)}$$

Applying these results to $f = \psi$ and simplifying yields:

$$p\epsilon'_\psi(p) = -\frac{(1 - \epsilon_\psi(p))\epsilon_\psi(p)}{1/\xi - 1}. \qquad \text{(B.13)}$$

Combining (B.10) with (B.13) gives

$$pF'(p; \kappa) = -\frac{p/\kappa}{(p/\kappa - 1)^2} + \frac{\epsilon_\psi(p)}{1/\xi - 1}\left(\frac{1}{1/\rho - 1} - \frac{1}{1/\xi - 1} \right)(1 - \epsilon_\psi(p)) \qquad \text{(B.14)}$$

Note that $F(p; \kappa) \geq 0$ is equivalent to:

$$\frac{1}{p/\kappa - 1} \geq \left(\frac{1}{1/\rho - 1} - \frac{1}{1/\xi - 1} \right)\epsilon_\psi(p) + \frac{1}{1/\xi - 1} \qquad \text{(B.15)}$$

and, upon simplification, $F'(p; \kappa) < 0$ is equivalent to

$$\frac{\epsilon_\psi(p)}{1/\xi - 1}\left(\frac{1}{1/\rho - 1} - \frac{1}{1/\xi - 1} \right)(1 - \epsilon_\psi(p)) < \left(\frac{1}{p/\kappa - 1} + 1 \right)\frac{1}{p/\kappa - 1}. \qquad \text{(B.16)}$$

In order to show that (B.15) implies (B.16), it will suffice to note that (B.15) implies

34

that $\frac{1}{p/\kappa-1} + 1 \geq \left(\frac{1}{1/\rho-1} - \frac{1}{1/\xi-1}\right)\epsilon_\psi(p)$, and so (B.16) is implied by

$$\frac{1}{1/\xi-1}(1 - \epsilon_\psi(p)) < \left(\frac{1}{1/\rho-1} - \frac{1}{1/\xi-1}\right)\epsilon_\psi(p) + \frac{1}{1/\xi-1}$$

which is true because it reduces to $0 < \epsilon_\psi(p)/(1/\rho - 1)$.

Step 3. Comparative Statics. The comparative statics with respect to κ follow from the first-order condition and the monotonicity properties of $\epsilon_\psi(p)$:

- For $\xi \in [-\infty, 0)$: $\epsilon_\psi(p)$ increases in κ

- For $\xi \in (0, 1)$: $\epsilon_\psi(p)$ decreases in κ

- For $\xi = 0$: $\epsilon_\psi(p)$ is constant in κ

This concludes the proof. $\qquad\square$

Proof of Proposition 3.10. Using the fact that $\psi(p) = \overline{\psi}p^\alpha$ for some $\overline{\psi} > 0$, we have $\psi^{-1}(x) = (x/\overline{\psi})^{1/\alpha}$ and so the equilibrium surplus in (2.16) ultimately simplifies to

$$S^*(\kappa, \lambda^*) = \rho\alpha \sup_{c>0} c^\rho - \kappa[\lambda^*\overline{\psi}/\rho]^{1/\alpha}c^{(1-1/\alpha)(\rho-1)+1}$$

while the social surplus in (2.17) simplifies to $\widetilde{S}(\kappa, \widetilde{\lambda}) = \sup_{c>0} c^\rho - \widetilde{\lambda}\overline{\psi}\kappa^\alpha c$. It follows that equilibrium and efficient consumption are $c^*(\kappa) = \overline{c}^*\kappa^{-\frac{\alpha}{1-\rho}}$ and $\widetilde{c}(\kappa) = \widetilde{\overline{c}}\kappa^{-\frac{\alpha}{1-\rho}}$, respectively, for some $\overline{c}^*, \widetilde{\overline{c}} > 0$. Proposition 3.5 shows that the equilibrium markup ratio is independent of κ and given by $m = 1 + (1/\rho - 1)/\alpha$, and so equation (2.12) implies that

$$\frac{1}{\mu^*} = 1 + \alpha(1/m - 1) = 1 - \frac{(1-\rho)\alpha}{\rho\alpha + 1 - \rho}. \tag{B.17}$$

Equation (2.20) then implies that the equilibrium and efficient M_e satisfy

$$\begin{aligned} M_e^*(fG(\overline{\kappa}^*) + f_e) &= \overline{T}(1 - 1/\mu^*) \\ \widetilde{M}_e(fG(\widetilde{\overline{\kappa}}) + f_e) &= \overline{T}(1 - \rho) \end{aligned} \tag{B.18}$$

and so combining (B.17), (B.18) with the fact that $\overline{\kappa}^* = \widetilde{\overline{\kappa}}$, we have

$$\frac{M_e^*}{\widetilde{M}_e} = \frac{1 - 1/\mu^*}{1 - \rho} = \frac{1/\rho}{m} < 1. \tag{B.19}$$

Now, the first equation in (B.18) and the resource constraint combine to give

$$M_e^* \int_{\underline{\kappa}}^{\overline{\kappa}} \phi(\kappa, p^*(\kappa)) \overline{c}^* \kappa^{-\frac{\alpha}{1-\rho}} G(\mathrm{d}\kappa) = \overline{T}/\mu^*.$$

When combined with the relationship

$$\frac{\phi(\kappa, p^*(\kappa))}{\psi(\kappa)} = \frac{\psi(p^*(\kappa))}{\psi(\kappa)} \times \frac{\phi(\kappa, p^*(\kappa))}{\psi(p^*(\kappa))} = \frac{m^\alpha}{\mu^*},$$

this implies that the equilibrium and equilibrium allocations satisfy

$$\widetilde{\overline{c}} \int_{\underline{\kappa}}^{\overline{\kappa}} \psi(\kappa) \kappa^{-\frac{\alpha}{1-\rho}} G(\mathrm{d}\kappa) = \frac{\rho \overline{T}}{\widetilde{M_e}} = m^{\alpha-1} \overline{c}^* \int_{\underline{\kappa}}^{\overline{\kappa}} \psi(\kappa) \kappa^{-\frac{\alpha}{1-\rho}} G(\mathrm{d}\kappa) \tag{B.20}$$

where we used the second equality in (B.19). With Cobb-Douglas technology, equilibrium allocations satisfy $p^* q^* = \alpha \psi(p^*) c^*$ and $n^* = (1-\alpha)\psi(p^*) c^*$ and efficient allocations satisfy $\kappa \widetilde{q} = \alpha \psi(\kappa) \widetilde{c}$ and $\widetilde{n} = (1-\alpha)\psi(\kappa)\widetilde{c}$. When combined with (B.20), this gives the result. $\qquad \square$

Proof of Lemma 4.2. Because $\hat{q}(p)$ is decreasing in p, the elasticity $\epsilon_u(q)$ will be decreasing in q if and only if $\ln \epsilon_u(\hat{q}(p))$ is increasing in p. We now define $W(p) :=$ $p[\ln \epsilon_u(\hat{q}(p))]' \int_p^\infty x\hat{q}'(x)dx$ and note that $W(p) \leq 0$ if and only if $[\ln \epsilon_u(\hat{q}(p))]' \geq 0$. Explicit calculation using (4.3) and $\int_p^\infty [x\hat{q}'(x) + \hat{q}(x)dx]dx = -p\hat{q}(p)$ then gives

$$W(p) = \left(1 + \epsilon_{\hat{q}}(p) + \frac{p^2 \hat{q}'(p)}{\int_p^\infty x\hat{q}'(x)dx}\right) \int_p^\infty x\hat{q}'(x)dx$$

$$= (1 + \epsilon_{\hat{q}}(p)) \int_p^\infty x\hat{q}'(x)dx + p^2\hat{q}'(p)$$

$$= -(1 + \epsilon_{\hat{q}}(p)) \int_p^\infty \hat{q}(x)dx - p\hat{q}(p).$$

Further, we have $W'(p) = -\epsilon_{\hat{q}}'(p) \int_p^\infty \hat{q}(x)dx$ and $\lim_{p\to\infty} W(p) = 0$, and so $W(p) \leq 0$ for all $p > 0$ if $\epsilon_{\hat{q}}'(p) \leq 0$ for all $p > 0$, which gives the result. $\qquad \square$