Equal Pay for Unequal Medicine*

Martin Schonger
ETH Zurich, mschonger@ethz.ch

February 2014

Abstract

Equal Pay for Unequal Medicine refers to the puzzle of patients receiving different qualities of a medical service even when, due to health coverage, they are charged the same price. The two proximate causes are: across providers quality varies even when prices do not, and, at any given provider patients receive different quality even when paying the same price. Moreover it has been observed that poor or minority patients are more likely to receive lower-quality care. While a positive correlation between income and quality holds in most markets, it is astonishing in this setting as price cannot link income with quality. Previous explanations of the disparities thus resort to additional assumptions like provider dislike of the poor, racism, exogenous geography, cultural misunderstandings or a conveniently signed correlation between income or race and patient preference for quality. This paper explains the fact pattern without such assumptions by the differential willingness of patients to incur acquisition costs. The model shows that even in the absence of patient heterogeneity beyond income, equal health coverage must be expected to lead to unequal treatment.

*This paper is based in parts on a chapter of my dissertation (Princeton, June 2012). I gratefully acknowledge comments on that chapter by Sylvain Chassang, Stephen Morris and Cecilia Rouse, as well as seminar participants at PET Taiwan and Zurich. All errors are mine. The paper is preliminary and I apologize to the reader for any unclarities and omissions, comments are most welcome.

1 Introduction

In its report "Unequal Treatment", the Institute of Medicine (Smedley et al., 2003) reviews a large number of empirical studies and concludes that poor and minority patients receive lower quality medical care than other patients even when paying identical amounts. Identical prices of medical services across hospitals and physicians for large groups of individuals are common due to third-party fee-for-service (FFS) health insurance. For example, most of Medicare is FFS and so regardless of patient income, race or ethnicity, or choice of provider, a specific medical service (i.e. billing code) costs the same[1] Medicare-mandated price. While in ordinary markets, a correlation between quality and socioeconomic indicators is explained by varying willingness to pay for higher quality, in this setting, where price is fixed, one might expect to find no large variations in quality correlated with patients' socioeconomic background. Attempts to attribute quality disparities to a potential correlation between patients' socioeconomic background and their medical condition have not been able to account for more than a fraction of the observed variation (Jha et al., 2007).

In addition to variation in quality by income among those who seek care, poor and minority patients seek less care even when on the same insurance plan, which has prompted Richman (2008) to ask whether on net insurance expansions redistribute money from poor to rich. This observation, that the uptake of almost free medical services is positively correlated with income is

[1] This assumes, as is the case for Medicare (see Shaviro 2004, p. 15), that patients cannot "top up", that is pay the physician or hospital an extra payment beyond the Medicare-specified fee.

not formalized in this paper, but could be formalized with a model similar to the ones proposed here and the result would obtain as long as there is at least some small inconvenience cost for taking up the "free" medical service.

Carlisle et al. (1997), using the California Hospital Discharge Data Set, examine 105,000 patients with a diagnosis indicating coronary heart disease. Controlling for age, sex, principal diagnostic code, the number of secondary diagnoses, elective, urgent or emergent admission, and insurance type they find that African American and Latino patients are much less likely to undergo three main cardiovascular procedures, namely coronary artery angiography, bypass graft surgery and coronary angioplasty. This discrepancy is manifest for patients whose insurance type is HMO, Medicare, Medicaid or who have no health insurance. Remarkably no such difference can be found comparing patients who are privately insured. The authors point out that cost-sharing under HMOs is typically minimal for these procedures, while private health insurance tends to have more cost sharing. Thus they conclude that the out-of-pocket expense borne by patients are an unlikely explanation for this pattern.

The literature has suggested many explanations for the quality disparity. These include differential physician perception of minorities,[2] residential segregation,[3] racial discordance between provider and patient,[4] different preferences of the poor,[5] low assertiveness, lack of information, the historic legacy

[2]E.g. van Ryn and Burke (2000), Abreu (1999), and Joe (1998).

[3]E.g. Chandra and Skinner (2003), Baicker et al. (2004) ,Baicker et al. (2005), and Wennberg et al. (2002).

[4]E.g. .Sohler et al. (2007) On the other hand, for a study that shows that African Americans receive less treatment regardless of their physician's race see Chen et al. (2001).

[5]E.g. Doescher et al. (2001), and Ayanian et al. (1999).

of segregated care,[6] provider prejudice,[7] and cultural misunderstandings.[8] The quality disparity can be attributed to two sources: (i) poorer patients frequent worse quality providers, and (ii) at any given provider poorer patients receive less quality. Estimates of how much of the quality disparity can be attributed to either of these two sources are rare. In the case of eye exams for diabetics Baicker et al. (2005) estimate that of the quality disparity found between the treatment black versus white patients receive, 56% is attributable to inter-provider variation, and 44% is due to intra-provider variation. Regardless of whether arising from across or within provider quality variation, in a setting where price is governmentally mandated and uniform across providers and patients, price cannot explain systematic quality variation.

This paper shows that the empirical pattern can be explained without recourse to any assumptions of patient heterogeneity beyond income. Some of the assumptions in the literature, could even be explained endogenously using models in the spirit of those put forward in this paper. The purpose is not to deny the existence of any of the factors the literature mentions, but rather to follow the scientific tradition of aiming for the simplest explanation possible. Thus the only heterogeneity between patients assumed here is patient income.

The intuition of the paper is that if there is some variable acquisition cost in addition to the uniform price, then differential patient willingness to incur these acquisition costs will lead to different outcomes even when paying the

[6]E.g. Smedley et al.(2003, p.103).

[7]E.g. Chandra and Staiger (2010). Schulman et al. (1999) is an audit study that reveals no concious or at least admitted bias, but does reveal bias in treatment recommendations

[8]E.g. Cooper and Roter (1998).

same. Under the term acquisition cost I subsume travel, search, bargaining, information gathering and similar costs. A fixed price restricts and distorts patients' choice, which they will partially undo by adjusting acquisition. If the restricted good is normal then the willingness to incur acquisition costs given a fixed price is increasing in income.

In the next section this insight is applied and demonstrated in three models that differ in the market structure and supply side. The first model addresses the case of the quality disparity arising from inter-provider variation. It demonstrates how poorer patients will on average end up with lower quality due to their selection of low-quality providers, even when they have the same preferences as their wealthier counterparts. The second and third models address the case of the quality disparity arising from intra-provider variation. Both models explain intra-provider variation by implicit bargaining over quality between patient and provider. Section 3 then instead of investigating the relationship in the context of specific market models, abstracts from these and posits that the quality a consumer receives is a function of the price she pays and how much of other goods she spends on acquisition. It is shown that not only in the case where price is unregulated, but even when government mandates a uniform price irrespective of income, quality will be increasing in income as long as it is a normal good. Section 3 concludes.

2 Models

In the following models I try to explain the empirically observed quality disparities as parsimoniously as possible and thus assume that patients are identical except for income. That assumption is not to be understood as an empirical claim, but rather aims to find out whether quality disparities can be explained solely on the basis of income. Therefore patients have identical preferences and the same medical condition, but there are two levels of income $j \in \{P, W\}$, poor patients with income ω^P and wealthy patients with income ω^W, where $\omega^W > \omega^P > 0$. There are two goods, a medical service with some non-negative scalar quality q, of which each patient consumes one unit and a numeraire good. All patients are on the same fee-for-service health plan, so physicians receive a fixed payment $\underline{p}$ for the medical service or procedure regardless of its quality. Physicians are prohibited from charging the patient in addition to that, and do not charge less as any savings would be kept by the health plan. Patients either have to pay nothing or a co-pay κ, so assume $0 \leq \kappa < \omega^P$. The co-pay is irrelevant for the results and just included so that it is clear that the results hold with or without a copay. Patients have strictly monotone, strictly convex preferences over quality and the numeraire good represented by a twice continuously differentiable utility function u. Their preferences are such that quality is a strictly normal good. Normality is used in the usual sense, that is, if patients could choose expenditure on the medical service, their chosen expenditure would be an increasing function of income. By definition strict normality of good 1 means that for all bundles $u_{12}u_2 - u_{22}u_1 > 0$ holds, where the subscripts denote partial derivatives.

2.1 Inter-provider disparity

Consider a simple travel cost model where two health care providers are located on opposite ends of a unit interval. There is a unit mass each of poor patients and wealthy patients. Both types of patients are uniformly distributed on the unit interval. Patients incur a travel cost τ per distance travelled, so a patient located at i $(0 \leq i \leq 1)$ incurs a travel cost τi to the left firm L, and $\tau(1-i)$ to the right firm R. Provider quality is exogenous and heterogeneous: Firm L offers quality q^L, firm R quality q^R. Without loss of generality assume $q^L < q^R$.

Proposition 1. ***[inter-provider variation]****Either all patients choose the high-quality provider, or the fraction of poor consumers choosing the low-quality provider is strictly higher than the fraction of rich consumers doing so.*

Proof. Denote by $\widetilde{i}^j$ the location of the left-most patient with income ω^j who purchases from firm R. Since firm R offers higher quality for the same co-pay it attracts more than half the patients, i.e. $\widetilde{i}^j \leq \frac{1}{2}$, and therefore it suffices to show that $\widetilde{i}^W = 0$ or $\widetilde{i}^W < \widetilde{i}^P$.

1. $\widetilde{i}^P = 0$: Then $\widetilde{i}^W = 0$, since by normality and the fact that a rich patient must be better off than a poor patient at the same location, $u\left(q^L, \omega^P - \kappa - \tau\widetilde{i}^P\right) \leq u\left(q^R, \omega^P - \kappa - \tau\left(1 - \widetilde{i}^P\right)\right)$ imply $u\left(q^L, \omega^W - \kappa - \tau\widetilde{i}^W\right) < u\left(q^R, \omega^W - \kappa - \tau\left(1 - \widetilde{i}^W\right)\right)$.

2. $\widetilde{i}^P \neq 0, \widetilde{i}^W = 0$.

3. $\widetilde{i}^P \neq 0, \widetilde{i}^W \neq 0$. Then both $\widetilde{i}^P, \widetilde{i}^W$ are indifferent between purchasing at either firm, i.e. for all j: $u\left(q^L, \omega^j - \kappa - \tau\widetilde{i}^j\right) = u\left(q^R, \omega^j - \kappa - \tau\left(1 - \widetilde{i}^j\right)\right)$.

Then either $\widetilde{i}^W < \widetilde{i}^P$, in which case we are done, or $\widetilde{i}^W \geq \widetilde{i}^P$. However, $\widetilde{i}^W$ is strictly better off than $\widetilde{i}^P$, and normality of quality implies that $\omega^W - \kappa - \tau\widetilde{i}^W - \left(\omega^W - \kappa - \tau\left(1 - \widetilde{i}^W\right)\right) > \omega^P - \kappa - \tau\widetilde{i}^P - \left(\omega^P - \kappa - \tau\left(1 - \widetilde{i}^P\right)\right)$, which implies that $\widetilde{i}^W < \widetilde{i}^P$. ■

The fraction of wealthy patients $1-\widetilde{i}^W$ choosing the higher quality physician is higher than the fraction $1 - \widetilde{i}^P$ of poor patients who do. Another way to look at this is to ask what fraction of the patients in the physician's waiting room are poor. At the low quality physician the fraction of poor patients is $\frac{\widetilde{i}^P}{\widetilde{i}^W+\widetilde{i}^P}$, which is larger than the fraction of poor patients $\frac{\widetilde{i}^P}{2-\widetilde{i}^W-\widetilde{i}^P}$ at the high quality physician.

Note that the self-sorting of patients by income occurs even though there is no correlation between patient income and patient location. If one were to endogenize location choice in this model residential segregation would endogenously emerge with poorer patients living near the low quality physician. Obviously there exists residential segregation for reasons unrelated to health care, the point here is that even in the absence of residential segregation by income the quality disparity should be expected to emerge.

2.2 Intra-provider variation

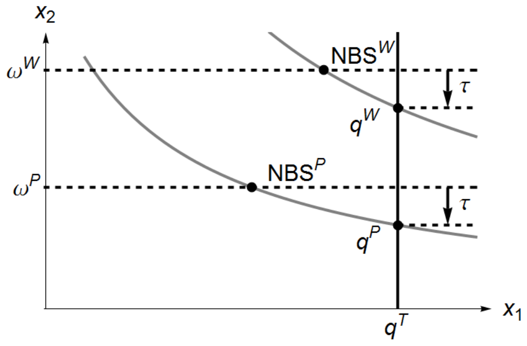
Even more puzzling than the phenomenon that poor patients frequent worse providers than their wealthy counterparts on exactly the same FFS-health coverage, is quality discrimination at a given physician. The physician receives the same payment from all patients, yet wealthy patients receive higher quality than poor patients. This observation can be explained by introducing

Nash bargaining between patient and physician. Rather than some haggling over quality in the practice, the bargaining should be imagined as implicit. Given FFS coverage, bargaining is only over quality, not also over price. Physicians have all the bargaining power.[9] The patients' outside option or threat point is to leave the physician's waiting room and seek care at another physician where they know they receive quality q^T for sure. But switching physicians is inconvenient and modelled by the switching cost $\tau > 0$. Therefore the patients' threat level of utility is $u\left(q^T, \omega^j - \kappa - \tau\right)$. Physicians observe patient income, this should not be taken literally when applying the model, but rather interpreted to mean that physicians observe some correlates of patient income and use those for statistical discrimination. The physician's threat point is to refuse treatment and to make zero profit. The profit function differs between the two following models that follow.

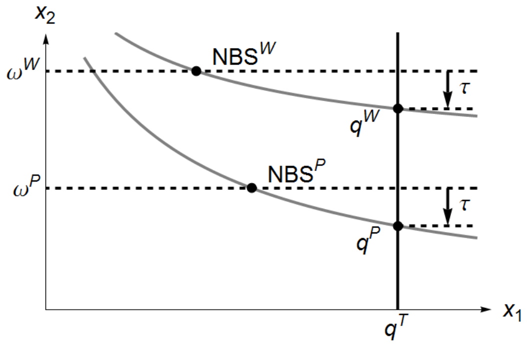
2.2.1 Intra-provider variation I: provider-effort

In the provider-effort model the physician can vary the amount of effort or time taken to perform the medical service. Quality is increasing in effort, but the effort is costly to the physician. This is captured by the fact that physician profit is decreasing in quality: $\pi(q) = \underline{p} - c(q)$. In order to have a bargaining problem there has to be a potential for bargaining surplus. Thus there must exist at least one type of patient who is strictly better off receiving the quality that makes the physician zero profit rather than

[9]If patients had all the bargaining power, then there would be no variable acquisition costs and thus all patients would get the same quality. There are no variable acquisition costs in the sense that in the Rubinstein game (Rubinstein, 1982), which provides a non-cooperative foundation for the Nash-Bargaining solution, the party with all the bargaining power is infinitely patient, that is, it has no costs from waiting.



(a) Normal good.



(b) Inferior good

Figure 1: Nash Bargaining Solution

the patient's threat point: there exists j such that $u\left(c^{-1}\left(\underline{p}\right), \omega^j - \kappa\right) > u\left(q^T, \omega^j - \kappa - \tau\right)$, where c^{-1} is the inverse of c.

Proposition 2. ***[intra-provider variation from provider effort]*** *Wealthy patients receive strictly higher quality than poor patients.*

Proof. Since the physician has all the bargaining power, in the Nash solution patients receive their threat level utility. Thus they are indifferent between staying or switching to another physician: $u\left(q, \omega^j - \kappa\right) = u\left(q^T, \omega^j - \kappa - \tau\right)$ for all j. For strictly monotone and convex preferences this equation implicitly defines q as a function of ω. That function is increasing in income ω if quality is a normal good as figure 1 illustrates: in (a) quality is a normal good, while in (b) it is an inferior good. ■

Even more surprisingly it is the poorer patient who receives the worse deal. Conventional price discrimination usually means that the poor get the better deal. Here price is administratively fixed by Medicare reimbursement rules, so medical providers can discriminate in quality alone. This quality discrimination has the opposite effect of the usual price discrimination in as much as it gives a worse deal to the poor.

2.2.2 Intra-provider variation II: choice of billing code

In the billing code model the physician cannot vary effort for a given medical service but can choose which of two medical services or procedures to administer. Both procedures are medically justifiable and get reimbursed. Procedure H has a higher quality than procedure L: $q^H > q^L$. The associated billing codes and thus reimbursements differ and are $\underline{p}^L$, $\underline{p}^H$. The physician incurs cost c^L, c^H for each procedure and his profit is the difference between the reimbursement and the cost. If procedure H is more profitable for the physician then there is no conflict of interest between the physician and the patient as both prefer procedure H which is therefore chosen. Therefore focus on the interesting case where there is a conflict of interest and assume that the profit of procedure L is larger than from procedure H: $\pi(L) = \underline{p}^L - c^L > \pi(H) = \underline{p}^H - c^H \geq 0$. Using the above framework the provider maximizes profit subject to the constraint that the patient receives no less than her threat utility $u\left(q^T, \omega^j - \kappa - \tau\right)$. Assume that it is feasible to do so, i.e. $u\left(q^L, \omega^j - \kappa\right) \geq u\left(q^T, \omega^j - \kappa - \tau\right)$.

Proposition 3. *[intra-provider variation from billing code]If* $u\left(q^L, \omega^W - \kappa - \tau\right) < u\left(q^T, \omega^W - \kappa - \tau\right)$ *and* $u\left(q^L, \omega^P - \kappa - \tau\right) \geq u\left(q^T, \omega^P - \kappa - \tau\right)$ *then wealthy patients receive high quality, poor patients the low quality procedure, else all patients receive the same quality procedure.*

Proof. (i) Suppose that the poorer patient undergoes the higher quality treatment H: Since the provider is foregoing the more profitable treatment L, we can conclude that L would lead to less than the poor patient's threat utility, i.e. $u\left(q^L, \omega^P - \kappa\right) < u\left(q^T, \omega^P - \kappa - \tau\right)$. But then normality of quality implies that $u\left(q^L, \omega^W - \kappa\right) < u\left(q^T, \omega^W - \kappa - \tau\right)$. Therefore if the poor patient receives the less profitable, higher quality treatment H, a fortiori the wealthy patient is also offered H.

(ii) Suppose that the richer patient is offered treatment L and chooses to undergo it: thus her utility under treatment L is at least as large as her threat utility, i.e. $u\left(q^L, \omega^W - \kappa\right) \geq u\left(q^T, \omega^W - \kappa - \tau\right)$. A fortiori by normality $u\left(q^L, \omega^P - \kappa\right) > u\left(q^T, \omega^P - \kappa - \tau\right)$, but then the provider will also only offer q^T to the poor patient who will accept.

(iii) The poorer patient is offered only treatment L, chooses to undergo it, while the wealthy patient is offered and accepts treatment H: The provider is foregoing the more profitable treatment L only for the wealthy patient, thus we conclude that L does not reach the threat utility of the wealthy patient, but does for the poor: $u\left(q^L, \omega^W - \kappa\right) < u\left(q^T, \omega^W - \kappa - \tau\right)$ and $u\left(q^L, \omega^P - \kappa\right) \geq u\left(q^T, \omega^P - \kappa - \tau\right)$. ∎

A model like this could explain findings like Baicker's (2004) observation that blacks have more money spent on them but receive the less effective treatments.

In both bargaining models in equilibrium poor and rich consumers spend the same on acquisition, that is neither incurs the switching cost τ. Their differential willingness to incur it is what drives the difference.

3 Acquisition Theorem

In the previous section acquisition activities were precisely modeled and it was shown that even for a fixed price the quality consumed is increasing in income if quality is a normal good. The question arises whether such results should be expected to hold under pretty much any setting with acquisition activities and quality being normal, or whether such extrapolations from the specific models are unwarranted. This section investigates that question. To do so acquisition is now modeled at a much more abstract but thus also more general level.

In models without acquisition activities the quality a consumer receives from a firm is a function of nothing more than the price P paid, that is quality x_1 is some function Q of P. To allow for acquisition activites such as transport, search, information gathering, or bargaining now I model quality x_1, as a function Q of not only the price P but also other inputs or goods. To do so consider a world with K goods. Good 1 is quality, the other goods can be consumed directly or be used to acquire good 1. Denote the amount of a good $k \geq 2$ consumed directly by x_k, and the amount used for acquisition by a_k. Quality is then a function of price P, and the amounts of the other goods used for acquisition: $x_1 = Q(P, a_2, .., a_K)$. Naturally assume that paying a higher price P or engaging in more acquisition increases quality, that is let Q be monotone. The appendix shows that under standard assumptions on Q as a production function there exists a demand function x^* and and acquisition function a^*. The demand function tells us given prices $p_2...p_K$ (there is no price given for good 1, as in lieu of such a price we have the acquisition function Q) and income ω how much of each good the individual consumes: $x^* : (p_2...p_K, \omega) \to (x_1^*, ..., x_K^*)$. In particular the demand function tells us how much quality the individual consumes, but it does not tell us how the consumer goes about getting that quality, that is what the acquisition function a^* tells us. Given prices $p_2...p_K$ and income ω it gives us how much the consumer will pay for good 1 and how much of the other goods she will use for acquisition: $a^* : (p_2...p_K, \omega) \to (P^*, a_2^*..., a_K^*)$. Thus so far price is not mandated by government and thus we would naturally expect, as we would in the absence of acquisition activities, that quality consumed is increasing in income. The following theorem states this unsurprising result:

Theorem 1. *If quality is a strictly normal good and price P is not regulated then,*

Quality is strictly increasing in income: $\frac{\partial x^*}{\partial \omega} > 0$.

The proof of this theorem and all following are in the appendix. Now consider the case where government mandates price for good 1. The *restricted* demand function tells us for prices $p_2...p_K$ and income ω how much of each good the individual consumes given some mandated price $\overline{P}$: $x^{**} : \left(p_2, ...p_K, \omega; \overline{P}\right) \to (x_1^{**}, ..., x_K^{**})$. The restricted acquisition function tells us how much the consumer spends on acquisition in addition to the mandated price $\overline{P}$ in order to get this quality x_1^{**}. $a^{**} : \left(p_2...p_K, \omega; \overline{P}\right) \to (a_2^{**}..., a_K^{**})$. The following theorem states that even when government restricts the price and forces consumers to spend $\overline{P}$ on good 1 irrespective of income, the quality consumed remains an increasing function of income:

Theorem 2. *If quality is a strictly normal good and price P is mandated to be equal to some $\overline{P} > 0$ then,*
Quality is strictly increasing in income: $\frac{\partial x^{**}}{\partial \omega} > 0$.

Thus even when government makes sure that all consumers pay the same price irrespective of income we should still expect the quality consumers receive to be increasing in income. Only in exceptional circumstances should this not hold. If competition is perfect and thus there is no room for acquisition activities to influence the quality one receives then such a government scheme should result in equal quality for all. But if there are acquisition activities then, as long as quality is a normal good, we should expect quality to still be increasing in income.

4 Conclusion

This paper develops a simple explanation for the empirically observed systematic quality disparity between patients of different socioeconomic backgrounds occurring even though price paid is the same for all patients due to fee-for-service (FFS) health coverage. While the empirical pattern of quality inequality would not be surprising in any market, conditional on price we would usually expect the disadvantaged to receive no less quality. The quality disparity arises from two factors, (i) poorer patients frequenting lower quality health care providers,[10] and (ii) poorer patients getting lower quality than their wealthier counterparts who frequent the same provider. To explain (i), inter-provider disparity a travel cost model is developed in this paper, while (ii), intra-provider disparity is explained using two Nash bargaining models. In all three models what drives the disparity in quality is the disparity in consumer income; consumers were assumed to be identical in all other aspects to show the strength of the theory. As price is fixed by government, the income disparity cannot produce the quality disparity via price. However, as long as quality of the medical service is not an inferior good, the richer a consumer, the higher is her willingness to engage in costly acquisition activities (e.g. search, travel, bargaining). In the travel cost model in equilibrium wealthier patients on average travel further than poor patients do. In the bargaining models in equilibrium no patient switches providers or spends any time on bargaining, but the differential willingness to do so leads wealthier patients to obtain higher quality from the same provider even as they pay the same, possibly zero, co-pay, and the provider gets the same payment from their health insurance.

By proposing this simple explanation for the empirically observed quality disparity and furthermore theoretically predicting that uniform restricted transfers, vouchers or FFS-health coverage will lead to systematic differences in quality received by income as long as there is some variable acquisition activity this paper advances the understanding beyond the current state of

[10] At the extreme a choice of lower quality could be interpreted as not seeing a health care provider at all, and thus the model can explain lower uptake as for example observed in Baicker (2004).

the literature as for example surveyed in Summers (1989), Bradford and Shaviro (2000), or Currie and Gahvari (2008).

The distinguishing feature of this paper's explanation of the quality disparity compared to prior explanations is its parsimony: while other explanations need to take recourse to additional assumptions of heterogeneity between patients such as residential location, health status or preferences, this paper explains the phenomenon by heterogeneity in income alone. This parsimony mirrors Becker (1968) who explains the higher incidence of property crime among the poor without postulating heterogeneous preferences or morals. Furthermore, like Becker this paper must not be understood to claim that there are no differences between poor and rich other than income, or that such differences could not explain different outcomes. Rather the point is to show that differential outcomes do not necessarily imply the existence and causality of other differences. Future empirical work is needed to distinguish between this parsimonious explanation and other factors on a case-by-case basis.

Similarly, correlations between race and outcomes, rather than income and outcomes can be explained, again like in Becker (1968), by the fact that even nowadays race remains a predictor of permanent income, even when controlling for transitory income. Conventional stories of different mores, genetically or culturally caused health conditions and behaviors, or widespread and shocking levels of racism among health care providers are, while in principle sufficient, again not necessary to explain such differences.

Arrow (1963) famously pointed to information asymmetries in medicine, explaining that understanding of treatment options and their quality is low among patients. If one interprets travel cost metaphorically as information acquisition cost then in this paper heterogenous and suboptimal levels of information arise endogenously in the travel-cost model, thus endogenizing Arrow's observation.

For policy this paper is highly relevant as it shows that the conventional wisdom that equalizing access to health care via uniform fee-for-service health coverage should lead to equal quality of treatment received is mistaken. Unequal quality for equal pay is not a puzzle, but follows from standard assumptions of consumer theory. On the contrary equal quality for equal pay for consumers with different incomes would be a puzzle. Such a puzzle could be explained by quality being a constant good,[11] quality not being variable to begin with, the absence of potentially variable acquisition costs or if acquisition requires spending resources that are more costly for the wealthy such as time.

Patients can learn from the observation of inverse quality discrimination that they are likely to get a better deal if they pretend to be rich rather than poor, which would be the opposite recommendation in the standard case of a market with price discrimination. Policymakers may want to rethink their preference for categorical equality. Fee-for-service health coverage in any case is only seemingly equal, but ends up giving more to the wealthy. If categorical equality in the quality of treatment is really desired then FFS would have to be adjusted such that provider reimbursements for each medical service are not uniform, but decreasing in patient income. As an alternative to

[11] As defined in chapter **??** constant good is a good for which spending is constant in income. For quasi-linear preferences the goods that are not quasi-linear are constant.

categorical equality which as this paper explains is costlier to achieve than one might think, policymakers could consider other options to improve the lives of the poor, as for example Deaton (2002) who calls for policymakers to "relax constraints on poor people tackling low incomes and poor education".

References

Abreu, J M, "Conscious and nonconscious African American stereotypes: impact on first impression and diagnostic ratings by therapists.," *Journal of Consulting and Clinical Psychology*, 1999, *67* (3), 387–93. 3

Arrow, Kenneth J., "Uncertainty and the welfare economics of medical care," *American Economic Review*, 1963, *53* (5), 941–73. 17

Ayanian, J Z, P D Cleary, J S Weissman, and A M Epstein, "The effect of patients' preferences on racial differences in access to renal transplantation," *New England Journal of Medicine*, 1999, *341* (22), 1661–9. 3

Baicker, Katherine, Amitabh Chandra, and Jonathan S Skinner, "Geographic variation in health care and the problem of measuring racial disparities," *Perspectives in Biology and Medicine*, 2005, *48* (1), S42–53. 3, 4

_ , _ , _ , and John E Wennberg, "Who you are and where you live: how race and geography affect the treatment of medicare beneficiaries," *Health Affairs*, 2004, *Web Exclusive*, 33–44. 3

Becker, Gary S., "Crime and punishment: an economic approach," *Journal of Political Economy*, 1968, *76* (2), 169–217. 17

Bradford, David F. and Daniel N. Shaviro, "The economics of vouchers," in C. E. Steuerle, ed., *Vouchers and the provision of public services*, Washington, DC: Brookings Institution Press, 2000, chapter 2, pp. 40–91. 17

Carlisle, David M., Barbara D. Leake, and Martin F. Shapiro, "Racial and Ethnic Disparities in the Use of Cardiovascular Prcedures: Associations with Type of Health Insurance," *American Journal of Public Health*, 1997, *87* (2), 263–7. 3

Chandra, Amitabh and Douglas O. Staiger, "Identifying provider prejudice in health care," Working Paper 16382, NBER 2010. 4

_ **and Jonathan S Skinner**, "Geography and racial health disparities," Working Paper 9513, NBER 2003. 3

Chen, Jersey, Siaf S. Rathore, Martha J. Radford, Yun Wang, and Harlan M. Krumholz, "Racial differences in the use of cardiac catherization after acute myocardial infarction," *New England Journal of Medicine*, 2001, *344* (19), 1443–9. 3

Cooper, L.A. and D.L. Roter, *Patient-provider communication: the effect of race and ethnicity on process and outcomes of healthcare*, Washington, D.C.: National Academies Press, 4

Currie, Janet and Firouz Gahvari, "Transfers in cash and in kind: theory meets the data," *Journal of Economic Literature*, 2008, *46* (2), 333–83. 17

Deaton, Angus, "Policy implications of the gradient of health and wealth.," *Health Affairs*, 2002, *21* (2), 13–30. 19

Doescher, Mark P., Barry G. Saver, Kevin Fiscella, and Peter Franks, "Racial/Ethnic inequities in continuity of care," *Health Services Research*, 2001, *36* (6), 78–89. 3

Jha, Ashish K, Douglas O Staiger, F Lee Lucas, and Amitabh Chandra, "Do race-specific models explain disparities in treatments after acute myocardial infarction?," *American Heart Journal*, 2007, *153* (5), 785–91. 2

Joe, JR, *The rationing of healthcare and health disparity for the American Indians/Alaska natives*, Washington, D.C.: National Academies Press, 3

Kannai, Y, "Concavifiability and constructions of concave utility functions," *Journal of Mathematical Economics*, 1977, *4* (1), 1–56. 24

Richman, Barak D., "Insurance Expansions: Do They Hurt Those They Are Designed to Help?," *Health Affairs*, 2008, *26*, 1345–57. 2

Rubinstein, Ariel, "Perfect equilibrium in a bargaining model," *Econometrica*, 1982, *50* (1), 97–109. 9

Schulman, Kevin A., Jesse A. Berlin, William Harless, Jon F. Kerner, Shyrl Sistrunk, Bernard J. Gersh, Ross Dube, Christopher K. Taleghani, Jennifer E. Burke, Sankey Williams, John M.

Eisenberg, and Jose J. Escarce, "The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization," *New England Journal of Medicine*, 1999, *340* (8), 618–26. 4

Shaviro, Daniel, *Who Should Pay for Medicare?*, University of Chicago Press, 2004. 2

Smedley, Brian D., Adrienne Y. Stith, and Alan R. Nelson, "Unequal treatment: confronting racial and ethnic disparities in health care," 2003. 2, 4

Sohler, Nancy, Lisa Fitzpatrick, Rebecca Lindsay, Kathryn Anastos, and Chinazo Cunningham, "Does patient-provider racial/ethnic concordance influence ratings of trust in people with HIV infection?," *AIDS and Behavior*, 2007, *11* (6), 884–96. 3

Summers, Lawrence H., "Some simple economics of mandated benefits," *American Economic Review*, 1989, *79* (2), 177–83. 17

van Ryn, M and J Burke, "The effect of patient race and socio-economic status on physicians' perceptions of patients.," *Social Science and Medicine*, 2000, *50* (6), 813–28. 3

Wennberg, John E, Elliott S Fisher, and Jonathan S Skinner, "Geography and the debate over Medicare reform," *Health Affairs*, 2002, *Web Exclusive*, W96–114. 3

5 Appendix

5.1 Preliminaries

Setup The consumer has an endowment $\omega > 0$, and thus her budget set is $\omega \geq p_2 x_2 + p_2 a_2 + P$ subject to $x_1 = Q(P, a_2)$.

Preferences are strictly monotone and strictly convex preferences over K goods which admit a twice continuously differentiable utility function u. Good 1 is the quality of the medical service. The medical service is an indivisible good and assume that it is given that the patient consumes the medical service. Assume that it is given that the patient consumes the medical service and the question is only at what quality. Goods 1 to K are strictly monotone. In addition good 1 (quality) is strictly normal.

The consumer has an endowment $\omega > 0$.

Goods $k \geq 2$ are sold on competitive markets at positive prices p_k. For convenience let 2 refer to 2 to K, so for example $p_2 = (p_2, .., p_K)$.

The quality a consumer receives is a function Q of the price P paid and how much of the other goods are spend on acquisition. Let Q be twice continuously differentiable function $Q = Q(P, a_2)$, where a_2 is short for the vector $(a_2, .., a_K)$. Let Q be strictly monotone in all inputs (P and each a_k). Let the production sets of Q be strictly convex and let Q exhibit decreasing returns to scale, i.e. assume that Q is strictly concave (. In addition assume that the price P is a strictly normal input.

Notation Subscripts denote (vectors of) partial derivatives, i.e. u_1 is the derivative of u with respect to x_1, u_x its gradient. Double subscripts denote second derivatives, accordingly u_{kl} is the scalar second derivative with respect to x_k, x_l, u_{xx} its $K \times K$- Hessian of u. By $\overrightarrow{1}_k$ vector of zeros with a 1 in the k. entry.

Technical assumptions To avoid dealing with non-negativity constraints, assume Inada-conditions on u and Q ensuring interior solutions.

Just to simplify the proofs and the already involved matrix algebra assume that a strictly concave representation of u has been chosen. Note that for bounded subsets a concave representation always exists if a quasi-concave one exists, see Kannai (1977).

To avoid non-substantial technical complications throughout the paper we shall abstract from the possibility of irregular points, meaning that a continuously differentiable, strictly monotone function never has a derivative equal to zero, i.e. for any differentiable function $f : \mathbb{R} \mapsto \mathbb{R}$ assume that If for all x,y such that $x < y$ then $f(x) < f(y)$, then $f_1 > 0$. This is a true restricting assumption, and not just choice of representation, but it is an issue only on a measure zero set and irrelevant for the economic substance of any of the results[12].

5.2 Proofs

Lemma 1. *Demand in the unrestricted problem exists.*

Proof. $\max_{x,P,a_2} u(x)$ s.t. $\omega \geq p_2^T x_2 + p_2^T a_2 + P$ and $Q(P, a_2) \geq x_1$. The objective function is (in particular) strictly quasi-concave. Though not linear any more

[12]An example of a function f where this technicality arises consider the simple scalar function $f : f(z) = z^3$. Its first derivative $3z^2$ is strictly positive for all z, except at 0.

as in the standard consumer problem, the feasible set is still weakly convex and compact. Therefore there exists a unique solution (x_2^*, P^*, a_2^*), which furthermore implies a unique $x_1^* \equiv Q(P^*, a_2^*)$. Uniqueness proves existence of both the demand function, $x^* = x^*(Q, p_2, \omega)$, and of the acquisition activity function, $a^* = a^*(Q, p_2, \omega)$. ■

Lemma 2.

$$\frac{\partial x^*}{\partial \omega} = \frac{\lambda^*}{\zeta u_2^T u_{22}^{-1} u_2 + \left(u_1 - u_{12}^T u_{22}^{-1} u_2\right)^2} \begin{pmatrix} u_1 - u_{12}^T u_{22}^{-1} u_2 \\ u_{22}^{-1}\left(\zeta u_2 - \left(u_1 - u_{12}^T u_{22}^{-1} u_2\right) u_{12}\right) \end{pmatrix}$$

$$\text{where } \zeta \equiv u_{11} - u_{12}^T u_{22}^{-1} u_{12} + \frac{u_1}{Q_a^T Q_{aa}^{-1} Q_a}$$

Proof. Under the Inada conditions any solution must be interior so we need not impose non-negativity constraints on x or a, and can write the patient's maximization problem as: $\max_{x,P,a_2} u(x)$ s.t. $\omega \geq p_2^T x_2 + p_2^T a_2 + P$ and $Q(P, a_2) \geq x_1$. Substituting $x_1 = Q(P, a)$, the Lagrangian with choice variables $x_2, a = (P, a_2)$ and Lagrange-multiplier $\lambda > 0$ (by strict monotonicity) can be written as: $\mathcal{L} = u(Q(P, a_2), x_2) + \lambda\left(\omega - p_2^T x_2 - p_2^T a_2 - P\right)$. Solution strategy: Given some x_1, we first define $(P^*, a_2^*) \equiv \text{argmin}\,(pa + P)$ s.t.$Q(P^*, a_2^*) \geq x_1$ and the acquisition cost function $A(x_1) \equiv \min_{P,a}(pa + P)$ s.t.$Q(P^*, a_2^*) \geq x_1$. x^* argmax $u(x)$ s.t. $\omega \geq px_2 + A(x_1)$.

Minimize total acquisition spending**:** $a^* = \text{argmin}\, A(P, a_2^*)$ s.t.$Q(P, a) \geq x_1$. $\mathcal{L} = px^A + P + \rho(x_1 - Q(P, a))$ (so $\rho > 0$). The first order conditions and the quality constraint can be written as $E(P^*, a_2^*, \rho^*; x_1) = \overrightarrow{0}$, where the function $E = E(P, a, \rho, x_1)$ is defined as:

$$E(P, a, \rho, x_1) = \begin{pmatrix} 1 - \rho Q_1 \\ p_2 - \rho Q_2 \\ x_1 - Q \end{pmatrix}$$

By the implicit function theorem:

$$\frac{\partial (P^*, a_2^*, \rho^*)}{\partial x_1} = -\left[E_{(a,\rho)}\right]^{-1} E_{x_1} = -\begin{pmatrix} \rho Q_{aa} & Q_a \\ Q_a^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} \overrightarrow{0} \\ 1 \end{pmatrix}$$

Which, using blockwise inversion[13] equals:

$$\frac{\partial (P^*, a_2^*, \rho^*)}{\partial x_1^*} = \frac{1}{Q_a^T Q_{aa}^{-1} Q_a} \begin{pmatrix} Q_{aa}^{-1} Q_a \\ -\rho \end{pmatrix}$$

Maximize utility:Substituting the acquisition cost function into the Lagrangian: $\mathcal{L} = u(x) + \lambda\left(\omega - p_2^T x_2 - A(x_1)\right)$. The first-order-conditions and the budget constraint can be written as $F(x^*, \lambda^*, \omega) = 0$, where the function $F =$

[13]For blockwise inversion note that in general for any non-singular square matrix M, any conform vector r, and any scalar α s.t. $\alpha - r^T M^{-1} r \neq 0$:

$$\begin{bmatrix} M & r \\ r^T & \alpha \end{bmatrix}^{-1} = \frac{1}{\alpha - r^T M^{-1} r} \begin{bmatrix} \left(\alpha - r^T M^{-1} r\right) M^{-1} + a^{-1} r r^T a^{-1} & -a^{-1} r \\ -\frac{1}{\alpha - r^T a^{-1}} r^T a^{-1} & 1 \end{bmatrix}$$

$F(x, \lambda, \omega)$ is defined as:

$$F(x, \lambda, \omega) \equiv \begin{pmatrix} u_1 - \lambda A_1 \\ u_2 - \lambda p_2 \\ \omega - p_2^T x_2 - A \end{pmatrix}$$

By the implicit function theorem:

$$\frac{\partial x^*, \lambda^*}{\partial \omega} = -\left[F_{(x,\lambda)}(x^*, \lambda^*, \omega)\right]^{-1} F_\omega(x^*, \lambda^*, \omega) = -\left[F_{(x,\lambda)}(x^*, \lambda^*, \omega)\right]^{-1} \mathbf{1}_{K+2}$$

The derivative of F w.r.t. the choice variables:

$$F_{(x,\lambda)}(x, \lambda, \omega) = \begin{pmatrix} u_{11} - \lambda A_{11} & u_{12}^T & -A_1 \\ u_{12} & u_{22} & -p_2 \\ -A_1 & -p_2^T & 0 \end{pmatrix}$$

From the cost minimization we know $A_1 = \rho^*$ and $A_{11} = -\frac{\rho}{Q_a^T Q_{aa}^{-1} Q_a}$.Evaluating at the optimum we have from the f.o.c.: $p_2 = \frac{u_2}{\lambda^*}$, $A_1 = \frac{u_1}{\lambda^*}$ and $\rho^* = \frac{1}{Q_1}$. Therefore:

$$F_{(x,\lambda)}(x^*, \lambda^*; \omega) = \begin{pmatrix} u_{11} + \frac{u_1}{Q_a^T Q_{aa}^{-1} Q_a} & u_{12}^T & -\frac{u_1}{\lambda^*} \\ u_{12} & u_{22} & -\frac{u_2}{\lambda^*} \\ -\frac{u_1}{\lambda^*} & -\frac{u_2^T}{\lambda^*} & 0 \end{pmatrix}$$

To find the inverse of $F_{(x,\lambda)}(x^*, \lambda^*; \omega)$ apply blockwise inversion twice, first finding the inverse of its following submatrix:

$$\begin{pmatrix} u_{11} + \frac{u_1}{Q_a^T Q_{aa}^{-1} Q_a} & u_{12}^T \\ u_{12} & u_{22} \end{pmatrix}^{-1} = \frac{1}{\zeta} \begin{pmatrix} 1 & -u_{12}^T u_{22}^{-1} \\ -u_{22}^{-1} u_{12} & \zeta u_{22}^{-1} + u_{22}^{-1} u_{12} u_{12}^T u_{2,2}^{-1} \end{pmatrix}$$

Applying blockwise inversion now to the entire matrix using the above inverse:

$$\frac{\partial x^*, \lambda^*}{\partial \omega} = \frac{\lambda^*}{\zeta u_2^T u_{22}^{-1} u_2 + \left(u_1 - u_{12}^T u_{22}^{-1} u_2\right)^2} \begin{pmatrix} u_1 - u_{12}^T u_{22}^{-1} u_2 \\ u_{22}^{-1} \left(\zeta u_2 - \left(u_1 - u_{12}^T u_{22}^{-1} u_2\right) u_{12}\right) \\ \zeta \lambda^* \end{pmatrix}$$

■

Lemma 3. $\frac{\partial x_1^*}{\partial \omega} > 0$.

Proof. As shown:

$$\frac{\partial x_1^*}{\partial \omega} = \lambda^* \frac{u_1 - u_{12}^T u_{22}^{-1} u_2}{\zeta u_2^T u_{22}^{-1} u_2 + \left(u_1 - u_{12}^T u_{22}^{-1} u_2\right)^2}, \text{ where } \zeta \equiv u_{11} - u_{12}^T u_{22}^{-1} u_{12} + \frac{u_1}{Q_a^T Q_{aa}^{-1} Q_a}$$

To shown that $\frac{\partial x_1^*}{\partial \omega} > 0$, note that the numerator is positive and $u_2^T u_{22}^{-1} u_2$ and ζ are both negative. The numerator is positive by normality of good 1. To see that $u_2^T u_{22}^{-1} u_2$ is negative, note that u_{xx} is negative, thus its submatrix u_{22} must be negative definite, and so is its inverse u_{22}^{-1} and by monotonicity u_2 is not the null vector. $Q_a^T Q_{aa}^{-1} Q_a$ is negative as Q_{aa} n.d. according to DRS and Q_a is not the null vector by monotonicity of Q. ■

Lemma 4. *Demand in the restricted problem exists.*

Proof. $\max_{x,a} u(x)$ s.t. $\omega \geq p_2^T x_2 + p_2^T a_2 + P$ and $Q(P, a_2) \geq x_1$ and $P = \overline{p}$.The objective function is strictly quasi-concave. The feasible set in this fixed-price problem is a subset of the feasible set in the unregulated price problem, it is still convex and compact. Therefore there exists a unique solution (x_2^{**}, a_2^{**}), which gives existence of the (on $P = \overline{p}$) conditional demand and acquisition activity functions, x^{**} and a_2^{**}. ■

Lemma 5.

$$\frac{\partial x^{**}}{\partial \omega} = \frac{\lambda^{**}}{\overline{\zeta} u_2^T u_{22}^{-1} u_2 + \left(u_{12}^T u_{22}^{-1} u_2 - u_1\right)^2} \begin{pmatrix} u_1 - u_{12}^T u_{22}^{-1} u_2 \\ \overline{\zeta} u_{22}^{-1} u_2 + \left(u_{12}^T u_{22}^{-1} u_2 - u_1\right) u_{22}^{-1} u_{12} \end{pmatrix}$$

$$\textit{where } \overline{\zeta} \equiv u_{11} - u_{12}^T u_{22}^{-1} u_{12} + \frac{{\lambda^{**}}^2 u_1}{u_2^T Q_{22}^{-1} u_2}$$

Proof. Again by the Inada conditions the solution must be interior, so the patient's maximization problem is: $\max_{x,a} u(x)$ s.t. $\omega \geq p_2^T x_2 + p_2^T a_2 + P$ and $Q(P, a_2) \geq x_1$ and $P = \overline{p}$.

Substituting $P = \overline{p}$ and $x_1 = Q(\overline{p}, a)$, the Lagrangian with choice variables x_2, a_2 and Lagrange-multiplier λ (by strict monotonicity) can be written as: $\mathcal{L} = u(Q(\overline{p}, a_2), x_2) + \lambda\left(\omega - p_2^T x_2 - p_2^T a_2 - \overline{p}\right)$. Solution strategy: For an arbitrary x_1 find $\overline{A}(x_1; \overline{p}) \equiv \min_{a_2^{**}} \left(\overline{p} + p_2^T a_2^{**}\right)$ s.t.$Q(\overline{p}, a_2^{**}) \geq x_1$. $\overline{A}(x_1; \overline{p})$ is the conditional acquisition cost function. $x^{**}(\omega) \equiv \operatorname{argmax}_{x^{**}} u(x)$ s.t. $\omega \geq p_2^T x_2 + \overline{A}(x_1; \overline{p})$. Then we can redefine a_2^{**} as a function of ω instead of $x_1 : a_2^{**}(x_1^{**}(\omega))$. Minimize indirect acquisition spending: $a_2^{**}(x_1) \equiv$ argmin pa s.t.$Q(\overline{p}, a_2^{**}) \geq x_1$. $\dfrac{\partial \mathcal{L}}{\partial a} = p_2^T - \rho Q_2 = 0$. The first order conditions and the quality constraint can be written as $\overline{E}(a_2^{**}, \rho^{**}, x_1) = 0$, where the function

$\overline{E} = \overline{E}(a, \rho, x_1)$ is defined as:

$$\overline{E}(a, \rho, x_1) = \begin{pmatrix} p_2^T - \rho Q_2 \\ x_1 - Q \end{pmatrix}$$

By the implicit function theorem:

$$\begin{pmatrix} \dfrac{\partial a_2^{**}}{\partial x_1} \\ \dfrac{\partial \rho}{\partial x_1} \end{pmatrix} = -\left(\overline{E}_{(a,\rho)}\right)^{-1} \overline{E}_{x_1} = -\begin{pmatrix} \rho Q_{22} & Q_2 \\ Q_2 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \overrightarrow{0} \\ 1 \end{pmatrix}$$

Which, using blockwise inversion to get $\left(F_{(a,\rho)}\right)^{-1}$ equals:

$$= \frac{1}{Q_2^T Q_{22}^{-1} Q_2} \begin{bmatrix} Q_{22}^{-1} Q_2 \\ -\rho \end{bmatrix}$$

Substituting the conditional (on $P = \overline{p}$) acquisition cost function into the Lagrangian: $\mathcal{L} = u(x) + \lambda\left(\omega - p_2^T x_2 - \overline{A}(x_1; \overline{p})\right)$. The first order conditions and the budget constraint can be written as $\overline{F}(x^{**}, \rho^{**}, x_1) = \overrightarrow{0}$, where the function $\overline{F} = \overline{F}(x, \lambda, \omega)$ is defined as:

$$\overline{F}(x, \lambda, \omega) \equiv \begin{pmatrix} u_1(x) - \lambda \overline{A}_1(x_1; \overline{p}) \\ u_2(x) - \lambda p_2 \\ \omega - p_2^T x_2 - \overline{A}(x_1; \overline{p}) \end{pmatrix}$$

By the implicit function theorem:

$$\frac{\partial (x^{**}, \lambda^{**})}{\partial \omega} = -\left[\overline{F}_{(x,\lambda)}(x^{**}, \lambda^{**}, \omega)\right]^{-1} \overline{F_\omega} = -\left[\overline{F}_{(x,\lambda)}(x^{**}, \lambda^{**}, \omega)\right]^{-1} \overrightarrow{1}_{K+2}$$

From the acquisition cost minimization, using first-order conditions and the envelope theorem we know $\overline{A}_1 = \rho^{**} =$, thus $\overline{A}_{11} = \rho_1^{**} = -\frac{\rho^{**}}{Q_2^T Q_{22}^{-1} Q_2}$. Also substitute the f.o.c. $\rho^{**} = \overline{A}_1 = \frac{u_1}{\lambda^{**}}$; $p_2 = \frac{u_2}{\lambda^{**}}$:

$$\overline{F}_{(x,\lambda)}\left(x^{**}, \lambda^{**}, \omega\right) = \begin{pmatrix} u_{11} + \frac{u_1}{Q_2^T Q_{22}^{-1} Q_2} & u_{12}^T & -\frac{u_1}{\lambda^{**}} \\ u_{12} & u_{22} & -\frac{u_2}{\lambda^{**}} \\ -\frac{u_1}{\lambda^{**}} & -\frac{u_2^T}{\lambda^{**}} & 0 \end{pmatrix}$$

Use blockwise inversion twice, first for submatrix :

$$\begin{pmatrix} u_{11} + \frac{u_1}{Q_2^T Q_{22}^{-1} Q_2} & u_{12}^T \\ u_{12} & u_{22} \end{pmatrix}^{-1} = \frac{1}{\overline{\zeta}} \begin{pmatrix} 1 & -u_{12}^T u_{22}^{-1} \\ -u_{22}^{-1} u_{12} & \overline{\zeta} u_{22}^{-1} + u_{22}^{-1} u_{12} u_{12}^T u_{22}^{-1} \end{pmatrix}$$

Applying block-wise inversion using the inverse of the submatrix from above, and substituting into the implicit function theorem:

$$\frac{\partial\left(x^{**}, \lambda^{**}\right)}{\partial \omega} = \frac{\lambda^{**}}{\overline{\zeta} u_2^T u_{22}^{-1} u_2 + \left(u_{12}^T u_{22}^{-1} u_2 - u_1\right)^2} \begin{pmatrix} u_1 - u_{12}^T u_{22}^{-1} u_2 \\ \overline{\zeta} u_{22}^{-1} u_2 + \left(u_{12}^T u_{22}^{-1} u_2 - u_1\right) u_{22}^{-1} u_{12} \\ \overline{\zeta} \lambda^{**} \end{pmatrix}$$

■

Lemma 6. $\frac{\partial x_1^{**}}{\partial \omega} > 0$.

Proof. As shown:

$$\frac{\partial x^{**}}{\partial \omega} = \frac{\lambda^{**}\left(u_1 - u_{12}^T u_{22}^{-1} u_2\right)}{\overline{\zeta} u_2^T u_{22}^{-1} u_2 + \left(u_{12}^T u_{22}^{-1} u_2 - u_1\right)^2}$$

$$\text{where } \overline{\zeta} \equiv u_{11} - u_{12}^T u_{22}^{-1} u_{12} + \frac{{\lambda^{**}}^2 u_1}{u_2^T Q_{22}^{-1} u_2}$$

Note that the numerator is positive and $u_2^T u_{22}^{-1} u_2$ and ζ are both negative. The numerator is positive by normality of good 1. To see that $u_2^T u_{22}^{-1} u_2$ is negative, note that u_{xx} is negative definite, thus its submatrix u_{22} must be negative definite, and so is its inverse u_{22}^{-1} and by monotonicity u_2 is not the null vector. $Q_a^T Q_{aa}^{-1} Q_a$ is negative as Q_{aa} n.d. according to DRS and Q_a is not the null vector by monotonicity of Q. $\blacksquare$