Job satisfaction and the wage gap.*

Eddie Dekel† — Northwestern and Tel Aviv Universities

Ady Pauzner — Tel Aviv University

October 20, 2014

Abstract

For many people there is tradeoff between choosing a job that they will enjoy and one at which they are good and will earn a high income. We embed this observation in a matching model. Consider then men and women who are a priori identical in the sense that both are equally likely to be good at one of two jobs and their satisfaction from each job is drawn from the same distribution. They are randomly matched into households after making a career choice, and have decreasing marginal utility of money. Thus, a career is chosen before knowing one's future spouse's income. If the distribution of enjoyment is log concave and single peaked, with the modal individual enjoying the job at which they are good, then there is either a unique symmetric equilibrium that is stable or an unstable symmetric equilibrium and two (mirror image) asymmetric equilibria that are stable. The latter display a wage gap and an opposite satisfaction gap, with one gender, wlog men, earning more even controlling for occupation. These equilibria display novel comparative statics. For example a tax on high wage couples results in women shifting into their more satisfying jobs and forgoing income (as one would expect), while interestingly men shift into higher income jobs, forgoing job satisfaction.

*We thank Gilat Levy, Yona Rubinstein, Annalia Schlosser, Balazs Szentes and Yoram Weiss for helpful comments. Dekel thanks the NSF (grant SES-1227434), and we both thank the Foerder Institute for Economic Research at Tel Aviv University, and the Pinhas Sapir Center for Economic Development at Tel Aviv University for financial support.

†Corresponding author. Telephone: +1-224-5222555. E-mail: dekel@northwestern.edu.

Keywords: Wage gap, job satisfaction, random matching, log-concave.

1 Introduction

When deciding on a career perhaps the most important questions a person considers are what she enjoys most and what she does best. While the occupation at which one has a comparative advantage – and hence that will generate higher income – is often the occupation one finds most satisfying, for many people these do not coincide. This paper is concerned with the tradeoff between job satisfaction and job suitability in occupation choice.

The optimal decision regarding this tradeoff depends on the income of one's spouse. The higher is this component of household income, the less important is additional income, and hence the more weight one will place on job satisfaction. As the choice of occupation is typically made before knowing the income of one's future spouse, it is the (equilibrium) distribution of incomes earned by the other gender that plays this role. This results in a game between the genders: how men trade off job satisfaction and suitability will affect the choice of women, and conversely.

We study how this feature of occupation choice may result in equilibria in which one gender, wlog males, emphasizes job suitability and the other gender, females, focuses on job satisfaction. In such an equilibrium females have, on average, lower suitability – and hence lower wages – in their chosen occupation. Thus, even in a symmetric environment where males and females choose each occupation in equal numbers and are a priori equally capable, female wages are lower. This wage gap continues to hold when controlling for occupation: in each occupation we observe males who chose it for the income and females who found it more satisfying. More generally, the model indicates that the job-satisfaction gap can explain part of the residual wage gap, and is consistent with the evidence that the wage gap favoring men coexists with a satisfaction gap that favors women.[1]

We also study the effect of changes in wages, divorce rates, and other parameters on these asymmetric equilibria. The results we obtain are of particular interest because they differ qualitatively from those of natural alternative models (such as those without households, e.g., where the wage gap is due to statistical or prejudicial discrimination). In addition to the direct effect – how one's choice responds to a parameter change when the other gender's behavior

[1] Of course there are other explanations for the job-satisfaction gap, such as selection and differential female participation in the work force.

is held constant – there is also an important indirect effect: how one's choice responds to the change in the other gender's behavior. We identify cases in which this opposing indirect effect dominates the direct effect; thus the comparative statics we obtain yield testable predictions and non-obvious policy implications.

Turning to the model's specifics, there is a continuum of individuals of each gender and there are two occupations. Individuals are characterized by two features: (1) the occupation at which they are most productive, which we call their *high-income occupation, HIO*; and (2) how much they enjoy the HIO relative to the alternative. Men and women are a priori identical, so their characteristics are drawn independently from a common distribution. We assume that the distribution of the job-satisfaction parameter is single peaked and log concave.[2] In addition, as individuals typically prefer the occupation at which they are better, we assume the modal type prefers their HIO. After choosing their occupation men and women are (randomly) matched into households. Individual utility is additively separable in job satisfaction and total household income and exhibits decreasing marginal utility in income.

Obviously any individual who enjoys his or her HIO more than the alternative will choose it. Indeed, the individual will continue to choose the HIO even when it is disliked, up to some threshold of dislike. Naturally, the optimal threshold for one gender is decreasing in the threshold of the other gender: if women, say, choose the HIO more often, men – who therefore expect higher income from their future spouse – would have lower marginal utility of further income and hence choose the HIO less often. Thus this is a game of strategic substitutes between the genders.

In this symmetric model there is always a symmetric equilibrium where males and females have the same threshold. This equilibrium may be (dynamically) stable or unstable depending on parameters. We show that if it is stable it is the unique equilibrium, while if it is unstable there is a unique asymmetric equilibrium (up to a relabeling of genders), which is stable. (Henceforth we refer to this pair of asymmetric equilibria that differ only in labels as the unique asymmetric equilibrium.) As discussed, in such an equilibrium one gender has higher income and the other more job satisfaction, and there is a wage gap controlling for occupation.[3]

[2] Most commonly studied distributions have log concave densities, see Bagnoli and Bergstrom and (2005).

[3] The model does not predict which gender will take each role; it only show that the asymmetric outcome

The result that for any parameters there is a unique stable equilibrium facilitates the study of comparative statics. Our focus in the study of comparative statics will be on the (stable) asymmetric equilibrium, as it constitutes our explanation of the wage gap and also displays the interesting testable implications mentioned above. We first obtain some general results on comparative statics in matching models such as ours, that we expect will be of use in other contexts, and then apply them to our specific model. These results show how and when information on the direction and relative sizes of the direct effects of parameter changes on the behavior of men and women identifies the direction of the equilibrium effect. As these direct effect are often easy to identify these results often generate clear-cut predictions. We now illustrate how these theorems can be applied in our context.

Consider for example a tax increase on high-income families (i.e., those comprised of two high-wage earners). The direct effect of such a tax increase is to reduce the incentives for both genders to choose the HIO. Indeed, we show that the effect on women (who choose the HIO less often than men) will be determined by this direct effect and they will further tend away from the HIO. However, the opposing indirect effect on men dominates the direct effect: the fact that women chose the HIO less often due to the tax increase leads men to choose it more often, even though the tax increase on its own reduces men's direct incentive to choose the HIO.

We also introduce a probability or time of being single. Comparative statics in this parameter could correspond to the changes in the wage gap due to the increase in divorce rates or in the age of marriage. Consider the case where parameters are such that the direct effect of an increase in the probability of being single will result in people choosing the HIO more often.[4] (This would be the case if an individual's marginal utility while single is larger than

---

can arise without assuming an a priori asymmetry between genders. We identify the high-wage gender with men, as is the case in reality. Explaining the selection between the equilibria (such as historical reasons) lies beyond the scope of our study. Nevertheless one can easily imagine that starting from a time when high income jobs required skills only available to men, then men would earn more and women would then tend to choose occupations based more on satisfaction than income, reinforcing men's choice of high income occupations, and so on, and that this would continue even as the high-income job only available to men would disappear. This in turn could raise the question of why men didn't shift the equilibrium to the opposite one (in the spirit of Doepke and Tertilt (2009) which they would prefer. This lies extremely far from our model, but in contrast to changing laws that give women the right to vote, changing the equilibrium requires solving a major coordination problem.

[4]The comparative statics for other cases can also be studied using our results and are noted in Section 4.2.

the marginal utility of money for that individual when he shares his money with a spouse.[5]) As before, the overall effect for women is the same as the direct effect. However, for men the indirect effect again dominates and they choose the HIO less often. Eventually, as the time or probability of being single increases further, the asymmetry (and wage gap) disappears, and the only equilibrium is the symmetric one (and further changes in the parameter of being single increase the choice of the HIO by both genders identically).

In addition, we briefly consider a variation of the model that highlights another aspect of the tradeoff between wages and job satisfaction. Individuals now have the same abilities in the two occupations and agree on which is more satisfying, and wages are decreasing in the number of people in each occupation. In equilibrium occupations become gender identified and have different wages: the occupation with greater job satisfaction is chosen more often by the population in general and relatively more often by one specific gender, and has lower wages.

Finally we note that beyond the substantive economics described above, this paper has two methodological contributions: The theoretical results on uniqueness of stable equilibria and those on comparative statics. These may be useful for other random-matching environments.

2 Related literature

There is a very large literature on the gender wage gap in earnings; for two excellent surveys see Altonji and Blank (1999) and Bertrand (2011). We discuss below a small subset of this literature, focusing on those theoretical models that are closest to our approach. The empirical literature, as far as we know, has not considered the connection between job satisfaction and the wage gap (with one exception, Zafar (2008), that we discuss subsequently). However, broadly speaking, these surveys document a wage gap in terms of men being more prevalent in better paying occupations, and in their receiving higher wages and promotions within

[5]This is more likely to be the case the smaller are household scale economies. (Since income is shared as a couple the benefit to an individual from higher income when with a spouse will tend to be smaller than when single; household economies may reverse this, but when they are small such a reversal will not occur.) Of course which marginal utility is higher depends on what individuals and couples consume, and many other features that we do not model (such as household bargaining).

occupations.[6] Moreover, the literature also documents a job-satisfaction gap (see, for example, Sousa-Poza and Sousa-Poza (2003)). We do not want to be interpreted as suggesting that the job-satisfaction gap is the whole story behind the wage gap; the literature offers many convincing explanations for large parts of the gap, ranging from human-capital investment, across psychological-behavioral reasons to statistical and taste discrimination. However, we do believe that job satisfaction plays an important role, and our model clarifies how this occurs and some of its implications.

Many papers provide theoretical explanations of the wage gap and gender specialization without assuming any direct taste for discrimination. One branch in this literature builds on firms that have incomplete information about employees. Lazear and Rosen (1990) is an influential example. In their model firms do not know how good each agent is at (future) household activities, and thus how likely it is that the agent will subsequently quit in favor of the alternative opportunity of working at home. Women are assumed to be (stochastically) better at household activities and thus more likely to quit. Since promotions require investments that will not be recouped if the worker quits, firms are less inclined to promote women.

A number of subsequent papers show, in models with this basic structure, that the wage gap and gender specialization arise even without assuming an a priori asymmetry between men and women. Instead, in these papers (as in ours) a critical role is played by the household which is comprised of a man and a woman who engage in complementary activities: earning income in the market or working at household chores. Firms have self-fulfilling expectations over the allocations of market and household tasks between men and women, generating an equilibrium with gender specialization and a wage gap. For example in Francois (1998) firms assign jobs that carry efficiency wages to men whose primary household role is to provide income and hence are less likely to shirk in favor of household demands. In Albanesi and Olivetti (2009) and Lommerud and Vagstad (2007) the driving force is that women, because they work more hours at home, have a higher cost of effort in market work. Hence it is more expensive to incentivize them to exert high effort (as in Albanesi and Olivetti), or – in the

[6] When controlling sufficiently narrowly for job categories – to the point that different levels of success can be considered different categories – the gap disappears; see Lazear and Rosen (1990).

absence of incentive contracts – it will not be worthwhile for the firm to invest in promoting women (as in Lommerud and Vagstad).

In the second branch of this literature (to which our paper belongs) firms have complete information about employees. Here, instead of incomplete information, the critical element is that individuals make career decisions before they are matched into a household. Again there are complementarities within the household as it needs both income and household production.

This literature began with the seminal work of Becker (1993). In his model households efficiently assign one spouse to invest in learning market skills and one to learn household related skills, and each will work in his or her area of expertise. Gender specialization – where women engage in one activity and men in the other – follows if there is even a slight comparative advantage for women in household in comparison to market activities.

Once again several papers show how gender specialization and a wage gap can arise without assuming an a priori asymmetry between men and women. For example, Hadfield (1999) shows that in a symmetric Becker-type model – one where the investment decision in learning household vs. market skills is taken before being randomly matched into households – such asymmetric equilibria exist. Engineer and Welling (1999) likewise obtain asymmetric equilibria in a symmetric model, but in addition expand the model to allow for heterogeneity in abilities, in which case there is also an equilibrium where individuals are trained according to abilities. More recently a number of papers, building on Peters and Siow (2002), consider pre-marital investments in models with various matching formulations and utility specifications (e.g., see Bhaskar and Hopkins (2011) and Booth and Coles (2009)). If, as seems reasonable, the choice of whether or not to invest is observable then after controlling for it there would be no wage gap. Our model obtains a gap controlling for all standard economic observables, such as investments and occupation, and calls for incorporating measures of job satisfaction as an additional control. There are many other papers which do not rely on incomplete information and study equilibria that display an asymmetry between men and women, but are even farther removed from our work.[7]

[7]For example, some papers focus on assortative matching and the competition between men or between women, or on inter-household bargaining. These include, among others, Iyigun and Walsh (2007), Cole, Mailath and Postlewaite (2001a, b, c), Echevarria and Merlo (1999), Felli and Roberts (2002), Ishida (2003),

As noted, our paper falls into this latter part of the literature where firms have complete information and males and females are a priori symmetric. Consistent with this literature we also obtain, under suitable parameters, asymmetric equilibria with gender specialization and a wage gap. Our model differs due to its focus on job satisfaction, the wage gap remaining when controlling for occupation and other economic observables, and the resulting comparative statics. As far as we know job satisfaction has not been studied as a possible component of the wage gap. One reason for this is surely that job satisfaction is more difficult to observe and hence include in empirical work as a control, but our result suggests the importance of attempting to do so. Finally, the paper also makes two methodological contributions. First, it shows that under weak assumptions on the distribution either the symmetric equilibrium is unique, or there is a unique pair of asymmetric equilibria that are stable (in addition to an unstable symmetric equilibrium).[8,9,10] Second, these assumptions generate the substantive comparative statics conclusions.

There is also a wide literature on job satisfaction, within which it has been argued that there is a job-satisfaction gap in favor of women (see, for example, Sousa-Poza and Sousa-Poza (2003)). While there may be many causes for this, it is consistent with our model of the wage gap in which women forgo income for job satisfaction, and men do the opposite.[11] Zafar (2008) is the only empirical paper that directly relates to our thesis. He studies the choices of major by college students, and – more importantly for our purposes – the reasons for their choices. Zafar argues that: "Males value pecuniary aspects of the workplace more, while females value non-pecuniary aspects of the workplace more", where non-pecuniary aspects

---

Peters and Siow (2002), Siow (1998), Elul, Silva-Reus and Volij (2002), and Nosaka (2007); see also Danziger and Katz (1996) for a cooperative perspective.

[8] In different models Bagnoli and Bergstrom (1993, 2005) show that log-concavity assumptions are useful for obtaining uniqueness of equilibrium.

[9] After completing this paper we became aware of Lommerud and Vagstad (2007) who argue that log-concavity yields these uniqueness properties of equilibrium. We do not see how their result could be correct as stated since additional assumptions – which in our model take the form of assuming that the distribution of how much individuals dislike the occupation at which they are better is single peaked where the modal type prefers the occupation at which they are better – seem necessary for the result.

[10] Bhaskar and Hopkins (2011) obtain unique equilibrim in an assortative matching model based on Peters and Siou (2002) by adding noise.

[11] More generally, Stevenson and Wolfers (2009) document a happiness gap in favor of women.

include, for example, "enjoy working at the jobs available after graduation."[12] These results are exactly in line with the predictions of our model.

3 The Model

There are two equally sized intervals of men (m) and women (w), and two occupations, A and B. Each person draws independently a type (k, x) where k is his/her high-income occupation (HIO) and x is his/her **dis**like of working at the HIO relative to the other occupation. An individual's HIO, k, is equally likely to be A or B and his/her relative dislike, $x \in \mathbf{R}$, is (independently) drawn according to a log-concave density f with support on an interval $[\underline{x}, \bar{x}]$, where f has a single peak below 0, and where we allow for $\underline{x} = -\infty$ or $\bar{x} = \infty$. An individual has income w_h from working in his/her HIO, and $w_l < w_h$ in the other profession. Thus the model is symmetric across men and women but there is heterogeneity in individuals' comparative advantage and in their preferences.

Individuals first choose a profession, and then some of them are randomly paired into households. Specifically, a proportion $1 - q$ of each gender is randomly matched and a proportion q remains single. The parameter q can be interpreted as a probability of being single or the expected proportion of life that individuals expect to be unmarried (either due to late marriage or divorce).[13]

The utility of agents is the sum of job-satisfaction utility, $-x$ or 0, and utility from household income. For a married individual the utility from income depends on the total income of the couple according to utility function u_C, while u_S denotes the utility function for money for single individuals. Thus the utility of an individual whose spouse earns w is:

	if single	if in a couple
choosing HIO	$u_S(w_h) - x$	$u_C(w_h + w) - x$
non-HIO	$u_S(w_l)$	$u_C(w_l + w)$

[12] His results are more subtle, comparing also workplace and non-workplace considerations, but in broad terms these are also consistent with our model.

[13] In reality changes in wages may affect q, but the effects of endogenizing q do not interact in interesting ways with our main concerns. We therefore chose to abstract from them and assume that q is exogenous.

We denote by U^H the increase in utility from the additional income due to choosing the HIO (ignoring job dissatisfaction, x) when the spouse has high income. Similarly U^L is this difference when the spouse has low income, and U^S is this difference when single. That is,

$$
\begin{aligned}
U^H &\equiv u_C(w_h + w_h) - u_C(w_l + w_h) \\
U^L &\equiv u_C(w_h + w_l) - u_C(w_l + w_l) \\
U^S &\equiv u_S(w_h) - u_S(w_l)
\end{aligned}
$$

We assume decreasing marginal utility of money: $U^L > U^H$ (and of course increasing utility in money: U^H, U^L and U^S are all strictly positive). We also want to avoid corner solutions where everyone chooses the HIO, hence we assume that the maximal dislike of an occupation, $\bar{x}$, is large enough to avoid this; a sufficient condition is $\bar{x} > qU^S + (1-q)U^L$. This implies that some individuals' dislike of the HIO is so high that for them it is dominant to choose the occupation that they enjoy more.

4 Equilibrium properties

This section contains important properties of the equilibria. These general results may be of use in other matching models as they rely on the structure of our model, but not its interpretation. These methodological results are used in the subsequent section to study the economic implications of the model. We start here by showing that either there is a unique equilibrium that is stable and symmetric or there is a unique pair of (mirror image) stable asymmetric equilibria and an unstable symmetric equilibrium. We then provide general results on how changes in parameters affect the equilibrium outcomes.

4.1 Characterization of the equilibria

Obviously an equilibrium has the form of a pair (x^m, x^w) of threshold strategies: men choose their HIO iff $x < x^m$, and women choose their HIO iff $x < x^w$. Recall that the larger is the threshold x^j of gender j, the more that gender gives up job satisfaction in favor of income. When randomly matched an individual of gender $j \in \{w, m\}$ meets a spouse with wage w_h with probability $F(x^{-j})$ and w_l otherwise (where $-j$ is the non-j gender). Thus the expected utility of an individual of gender j is

$$(1-q)\left[F\left(x^{-j}\right)u_C\left(w_h+w_h\right)+\left(1-F\left(x^{-j}\right)\right)u_C\left(w_h+w_l\right)\right]+qu_S\left(w_h\right)-x$$

if the individual chooses the HIO, and

$$(1-q)\left[F\left(x^{-j}\right)u_C\left(w_l+w_h\right)+\left(1-F\left(x^{-j}\right)\right)u_C\left(w_l+w_l\right)\right]+qu_S\left(w_l\right)$$

otherwise.

Gender j's best-reply threshold, x^j, given the opposite gender's threshold, x^{-j}, is then

$$x^j=B\left(x^{-j}\right)\equiv q\cdot U^S+(1-q)\cdot\left[U^H F\left(x^{-j}\right)+U^L\left(1-F\left(x^{-j}\right)\right)\right].$$

Since $U^L>U^H$ we see immediately that the slope of the best-reply function is negative: if one gender chooses the HIO more often then the other gender chooses it less. (If x^j increases and $U^L>U^H$ then $U^H F\left(x^j\right)+U^L\left(1-F\left(x^j\right)\right)$ decreases.)

A pair of thresholds (x^m,x^w) is then an equilibrium if $x^m=B\left(x^w\right)$ and $x^w=B\left(x^m\right)$. (Note that $B\left(x^j\right)$ is the best reply function of gender $-j$, not j.) In general there can be two types of equilibria: (1) symmetric, in which $x^m=x^w$, where we will denote the common equilibrium threshold by x^s; and (2) mirror-image asymmetric equilibria, in which case we focus throughout, wlog, on the equilibrium with $x^m>x^w$.

We are interested in (dynamically) locally stable equilibria. An equilibrium is stable in this sense if, starting from near enough to an equilibrium, the behavior would converge back to the equilibrium, where the dynamics are given by the best-response functions. An equilibrium is unstable if it locally diverges. It is straightforward that an equilibrium (x,y) is stable if $B'(x)\times B'(y)<1$ and it is unstable if $B'(x)\times B'(y)>1$. In general if $B'(x)\times B'(y)=1$ an equilibrium may be neither stable nor unstable, but we will see that in our model such equilibria are stable.

Proposition 1 *Depending on the model's parameters (q,u_C and u_S), either there is a unique equilibrium x^s which is stable and symmetric with $\left|B'\left(x^s\right)\right|\leq 1$, or there are three equilibria: an unstable symmetric equilibrium x^s with $\left|B'\left(x^s\right)\right|>1$ and two stable asymmetric equilibria (x,y) and (y,x) with $B'(x)\times B'(y)<1$.*

Proof: See appendix.

4.2 Comparative statics

The comparative statics results obviously depend on two effects. First, there are the standard direct effects: how each gender's choices respond to a parameter change when the other gender's behavior is held constant. Second, there are the indirect effects: each gender's behavior does change, which further impacts the other gender's choices. The results in this section show how the overall equilibrium effect can be determined from the direct effects alone.

To state these results formally let t be an exogenous parameter affecting both genders, with $t = 0$ denoting the initial situation. In this subsection we thus add the argument t to all functions. So $x^s(t)$ denotes the symmetric equilibrium as a function of t, that is, $x^s(t) = B(x^s(t), t)$. Similarly, an asymmetric equilibrium is a pair $(x^m(t), x^w(t))$ that solves $x^j(t) = B(x^{-j}(t), t)$ for $j = m, w$. Denote partial derivatives using subscripts, for example $B_t(x^s(t), t) = \partial B(y, t)/\partial t$ at the point $y = x^s(t)$.

Theorem 1 states that in the case of a (stable) symmetric equilibrium the combined equilibrium effect turns out to be of the same sign as the direct effect.

Theorem 1 *Consider a stable symmetric equilibrium $x^s(t)$. Then at $t = 0$, $x_t^s(t)$ has the same sign as $B_t(x^s(t), t)$.*

Theorem 2 considers (stable) asymmetric equilibria. In this case the relationship depends on the signs of the direct effects and their relative magnitudes. If the direct effects on the two genders go in opposite directions (part 1) then the combined equilibrium effect is the same as the direct effect for each, and, moreover the effect on men is larger. If the direct effects are in the same direction, there are two cases: If the direct effect on women is larger than that on men (part 2) the combined effect on women is the same as the direct effect, while the combined effect on men is the opposite. Otherwise (part 3) at least one of the combined effects must be the same as the direct effect. However, which of the three possibilities – whether $x^m(t)$ or $x^w(t)$ or both – change in the same direction as the direct effect cannot be determined without further data.

Theorem 2 *Consider a stable asymmetric equilibrium $(x^m(t), x^w(t))$, with the convention that $x^m > x^w$. Then at $t = 0$:*

1. *If* $B_t(x^j(t), t) > 0 > B_t(x^{-j}(t), t)$ *for* $j = m$ *or* w,[14] *then*

$$x_t^j(t) > 0 > x_t^{-j}(t).$$

Moreover $|x_t^m(t)| > |x_t^w(t)|$.

2. *If* $|B_t(x^m(t), t)| \geq |B_t(x^w(t), t)| > 0$, *then*

$$\mathit{sign}(x_t^w(t)) = \mathit{sign}(B_t(x^m(t), t)) \text{ and } \mathit{sign}(x_t^m(t)) = -\mathit{sign}(B_t(x^w(t), t)).$$

3. *If* $0 < |B_t(x^m(t), t)| < |B_t(x^w(t), t)|$ *then*

$$\mathit{sign}(x_t^w(t)) = \mathit{sign}(B_t(x^m(t), t)) \text{ or } \mathit{sign}(x_t^m(t)) = \mathit{sign}(B_t(x^w(t), t)).$$

These theorems follow, with elementary algebraic manipulations, from the next two lemmas.

Lemma 1 *In a stable asymmetric equilibrium* $(x^m(t), x^w(t))$, *with the convention that* $x^m > x^w$, *at* $t = 0$,

$$|B_x(x^m(t), t)| < 1 < |B_x(x^w(t), t)|.$$

Proof. At $t = 0$, $B_x(x^j(t), t) = (1 - q)\left(U^H - U^L\right) f(x^j)$ so $|B_x(x^m(t), t)| < |B_x(x^s(t), t)| < |B_x(x^w(t), t)|$ where x^s denotes the unstable symmetric equilibrium. Recall that $x^m > x^s > x^w$, and that f is decreasing in this region. Since $|B_x(x^s(t), t)| > 1$ (by instability) and $|B_x(x^m(t), t)| |B_x(x^w(t), t)| \leq 1$ (by stability) we have

$$|B_x(x^m(t), t)| < 1 < |B_x(x^w(t), t)|.$$

■

Lemma 2 *In a stable equilibrium* $(x^m(t), x^w(t))$, *at* $t = 0$,

$$\mathit{sign}(x_t^m(t)) = \mathit{sign}(B_t(x^w(t), t) + B_x(x^w(t), t) B_t(x^m(t), t))$$

and likewise

$$\mathit{sign}(x_t^w(t)) = \mathit{sign}(B_t(x^m(t), t) + B_x(x^m(t), t) B_t(x^w(t), t))$$

Proof. See appendix. ■

[14] *Here one inequality may be weak.*

5 The effect of exogenous changes in parameters

Would a tax reduction on high wages lead to an increase in the choice of the HIO? Is such a change equivalent to a tax increase on low wages? What would lead to a reduction in the asymmetry between the genders in the choice of HIO, and hence in the wage and job-happiness gaps? What are the model's predictions regarding the effects of the documented increase in the time spent single or of reforms in the laws regarding post-divorce sharing of income?

In this section we apply our general results of the preceding section to answer such questions. The comparative-statics results we obtain are also useful for identifying the testable implications of this model. Since our focus throughout is on the asymmetric stable equilibrium we only describe the comparative statics for this equilibrium.

As discussed the overall effect of changing such parameters is determined by the direct and indirect effects. Here we further decompose the direct effect so that the overall effect can be better understood by considering three ingredients. First, there is the direct incentive effect: a change in a gender's incentive to select the HIO holding fixed the behavior of the other gender and the income of the other gender. Second, there is the direct wealth effect: the change in the income provided by one's spouse holding fixed their action. Finally, there is the indirect effect, resulting from changes in the behavior of the other gender.

To illustrate these effects we consider three changes that, at first glance, could be expected to increase the choice of the HIO by both genders. These changes are: (1) an increase in w_h, (2) a decrease in w_l, and (3) an increase in $U^H = u_C(w_h + w_h) - u_C(w_h + w_l)$, (holding constant U^S and U^L). As noted, the latter could arise from a reduction in taxes on households with two high incomes.

Holding all else fixed, an increase in the wages one expects to get from the HIO, w_h, will increase the incentive to choose the HIO. However, an increase in w_h also increases the expected income of one's (future) spouse, and this wealth effect works in the opposite direction. Indeed the overall direct effect cannot be signed, hence neither can the indirect effect be signed. Hence we cannot say anything about the effects of such a change.

On the other hand, a decrease in w_l has a clear direct effect: the lower income of the non HIO, and the lower expected spousal income, both increase the direct incentive to choose the HIO. This implies that, if we are in a symmetric stable equilibrium, a small decrease in w_l

will increase the (common) equilibrium threshold below which the HIO is adopted. However, in a stable asymmetric equilibrium the threshold for males or females can go up or down because the equilibrium effect (of the other gender choosing the HIO more often) may or may not dominate the direct effect. Of course, it cannot be that both genders choose the HIO less often.

Interestingly an increase in U^H has an unambiguous (and perhaps surprising) effect on male and female equilibrium thresholds. When the stable equilibrium is asymmetric, the female threshold unambiguously increases and the male threshold decreases. This is because for males the indirect effect – of females choosing the HIO more often – must dominate the direct effect.

Why is the comparative static on U^H unambiguous? Details follow from the proofs of the results of Section 4.2, but we provide the basic ideas here. Recall that U^H is the utility differential from bringing wage w_h vs. w_l given that the spouse brings wage w_h. The direct effect of a change in U^H is stronger for females, who are more likely than men to face a spouse who earns w_h. Moreover, we show that males react to the change in the females' threshold more strongly than the change in the females' threshold itself (i.e., the slope of males' threshold as a function of females' threshold is steeper than 1). Combining these two arguments implies that the indirect effect on men dominates the direct effect on them. Thus, the overall effect on males must be a decrease in their threshold. For females the opposite holds since the slope of the females' best-reply function is less than 1 and the males' direct effect is smaller than that of females.

Perhaps even more intriguing than the effect of changing U^H is the effect of a change in divorce laws that increases post-divorce income sharing. Specifically assume that after divorce each individual gives a portion $\alpha \in [0, 1/2]$ of their income to their ex.[15] The direct effect of increasing α is to decrease the incentive to choose the HIO. However, as in the case of U^H, the direct effect on women is stronger than on men. Thus, as one would expect, some women keep less of the post-divorce income, they decrease their choice of the HIO. However, for men the indirect effect dominates and they *increase*, rather than decrease, their choice of

[15] Literally speaking the parameter α is not in our formal model presented above, but it is obvious how to include it and we provide a formal model with this sharing rule in the proofs that are in the appendix.

the HIO, giving up job satisfaction and for a higher paying occupation when the share of their post-divorce income decreases.

We summarize the effect of exogenous changes in some parameters that are common to both genders in the table below. We consider changes in the wages w_l and w_h, and also changes in U^H, U^L and U^S which correspond, for example, to changes in taxes on high income couples (those were both individuals earn w_h), low income couples (where both earn w_l), or on income of singles. We also consider changes in q, the expected proportion of time spent single, and in α, the proportion of income that continues to be shared after divorce. (The proofs for the claims in the table are in the appendix.)

We find the cases where the effects on men and women go in opposite directions to be especially interesting. These results are a consequence of the interplay between the genders' career choices that appear in our model. They would not arise in a model where only direct effects are considered, for example if the wage gap is explained by differences between the genders or in how they are treated and the interdependence of their choices is ignored.

change in	changes in the equilibrium thresholds $(x^m(t), x^w(t))$
$U^S \uparrow$	$(\downarrow, \uparrow)$
$U^H \uparrow$	$(\downarrow, \uparrow)$
$U^L \uparrow$	not $(\downarrow, \downarrow)$
$q \uparrow$[16]	$\begin{cases} \text{if } U^S > U^H & (\downarrow, \uparrow) \\ \text{if } U^S < U^L & \text{not } (\uparrow, \uparrow) \end{cases}$
$w_l \uparrow$	not $(\uparrow, \uparrow)$
$w_h \uparrow$	depends on u_C''' [17]
post-divorce share $\alpha \uparrow$	$(\uparrow, \downarrow)$

[16] The row corresponding to q differs in that it has two ranges. The effect of an increase in q depends on whether it increases the direct effect on women of choosing the HIO more often. This in turn depends on whether additional income for women is more important when single or in a couple. Formally, this corresponds to how U^S compares to $U^w \equiv U^H F(x^m) + U^L (1 - F(x^m))$. While this comparison depends on the equilibrium threshold levels, a sufficient condition for $U^S > U^w$ is that $U^S > U^L$, and similarly a sufficient condition for $U^S < U^w$ is $U^S < U^H$. Whether increases in income matter more to an agent who is single or married will depend on factors such as the extent to which consumption in couples is of private or public goods, and how usage and needs for money differ between singles and couples.

[17] Increases in w^h will increase the incentive to chose the HIO if one is single or one's spouse brings low

6 Concluding remarks

In this section we first consider a different model: where one occupation is more enjoyable for everyone. Then we return to the main model of the paper and consider various issues and extensions. These include whether job satisfaction is private or public, welfare comparisons, and how the predictions of our model relate to predictions of alternative models.

6.1 Enjoyable occupations

In the previous sections we explored a model that is symmetric both in terms of males and females and in terms of the occupations. We identified an equilibrium where those of one gender tend to choose their preferred job and those of the other gender select the occupation with higher income. The model generates a wage gap even controlling for occupation.

However, it seems that some occupations are chosen primarily by females and have lower income. Becker's seminal paper can immediately explain this for occupations that facilitate domestic activities, for example, occupations that provide flexibility in caring for children. Yet certain lower-income occupations, such as acting or nursing, that do not appear to be particularly suited for spouses responsible for domestic activities, seem to attract females more than males. Can such choices be explained without assuming an asymmetry between men and women?

We now present a variation of our preceding model that does so. It builds on the same main features as before, namely the tradeoff between income and job satisfaction, taking into account that household income will include that of one's (future) spouse, and continuing to assume decreasing marginal utility of household income. We make two main changes. First, all individual agree that one specific occupation is more satisfying than the other. Second (so that markets will clear given this common preference), wages in an occupation are decreasing in the number of individuals choosing that occupation. We are interested in equilibria where one gender chooses primarily a more satisfying but lower paying occupation and the other primarily chooses the occupation where higher wages compensate for lower satisfaction. In

income, but if the spouse brings high income then the value of additional high income (the benefit of changing the couples income from $w_l + w_h + t$ to $2\,(w_h + t)$) depends on the rate at which marginal utility is decreasing, i.e., on u_C'''.

contrast to the model analyzed in the preceding sections, the wage gap generated by the alternative model of the current subsection would disappear when controlling for occupation. Instead, the current model obtains gender-identified occupations that differ in the proportions of females to males.

Specifically, we assume that there are two occupations $i \in \{A, B\}$ with wages decreasing in the number of individuals, n_i, choosing an occupation. While the function determining wages, $w(n_i)$, is the same in both occupations, A is liked more by everyone by a fixed amount $x > 0$. As before there is a continuum of mass 1 of individuals of each gender.

Thus, utility from working in A when married to a spouse whose income is w equals $u(w(n_A) + w) + x$, while the utility from B is $u(w(n_B) + w)$, where as before $u' > 0$ and $u'' < 0$. After their choice of occupation agents from different genders are randomly matched. Again, as before there are two types of equilibria, symmetric and asymmetric, and we focus on the asymmetric equilibrium which has a wage gap. In such an equilibrium all individuals of one gender, wlog women, choose A and men split between the occupations so that they are indifferent:

$$\begin{aligned} u(w(n_A) + w(n_A)) + x &= u(w(n_B) + w(n_A)), \text{ i.e.,} \\ u(2w(n_A)) + x &= u(w(2 - n_A) + w(n_A)). \end{aligned}$$

(This equation has a solution as long as x is not too large.) In this equilibrium more individuals choose A (all women and some men) than B (the remaining men), men are indifferent between the two occupations, and women strictly prefer the choice of A.

The preferred occupation naturally attracts more workers and has lower wages. In equilibrium it becomes identified with one gender that then has lower wages even though both occupations have the same wage function $w(\cdot)$. Note that women prefer this equilibrium, while men prefer the mirror-image equilibrium. (Of course women would prefer even more receiving the high wages in the more satisfying occupation, but that cannot occur in equilibrium.)

6.2 Private or public benefits from job satisfaction

In our model a critical feature is that the income an agent brings to the household is a pubic good – i.e., both spouses enjoy it – with decreasing returns. For notational simplicity the model was written as if job satisfaction affects only the individual choosing the job. However, none of our results would change in any way if job satisfaction were also a public good, i.e., if each person benefited when his or her spouse chose an enjoyable job. More precisely, if a spouse's job dissatisfaction, x, also enters as an additive term in one's utility, then all the equations determining individual decisions are unaffected. Indeed the equilibrium is unchanged. This holds even if an individual's enjoyment of a spouse's job satisfaction is partial. Thus the critical feature of job satisfaction that is necessary for our conclusions is that – unlike income – job satisfaction does not have decreasing marginal utility: one's enjoyment of one's spouse's job satisfaction does not decrease in one's own job satisfaction. The assumption of whether the benefits are fully or partially public or private per se is of no consequence for all the preceding analysis.

6.3 Welfare

There are two potential sources of inefficiency in our main model. First, individuals may fail to take into account the benefit income brings to their spouse. This is a standard externality issue. Second, ex ante, due to decreasing marginal utility of income, it is socially better to match a high-wage male with a low-wage female, and a high-wage female to a low-wage male, than to match the two high-wage earners together and the two low-wage earners together. That is, negative assortative matching would be efficiency enhancing.

The first potential source of inefficiency does not exist if job satisfaction is public (in the sense discussed in subsection 6.2). This is because then the individual tradeoff between income and job satisfaction is the same as the couple's welfare-maximizing tradeoff.

Now consider the second potential source of inefficiency, that making the matching more negatively assortative may enhance welfare. To focus on this inefficiency we first consider the case where job satisfaction is public (and hence where the preceding inefficiency does not exist). In this case, surprisingly, it turns out that the stable equilibrium is ex ante constrained efficient in the sense that no other thresholds for males and females can improve ex ante expected payoffs. This is the appropriate notion of efficiency if we only permit interventions that change agents decisions, i.e., their thresholds. We thus allow for interventions that force – or incentivize – individuals to choose different thresholds, but not interventions that change the exogenous matching. (Constrained efficiency *does* consider long-run effects of interventions; it just restricts those interventions to take the random matching process as given.)

We now sketch the main argument that the equilibria are constrained efficient. First note that given any pair of thresholds, if one gender moves in the direction of improving its interim payoff (i.e., certain types of dislike of men, say, choose a different action) then the other gender gains ex ante as well.[18] Thus the ex ante efficient thresholds constitute an equilibrium. If the symmetric equilibrium is stable then it is unique, hence it must be ex ante constrained efficient. If the symmetric equilibrium is not stable we now argue why it cannot ex ante dominate the asymmetric equilibrium, and hence that the stable asymmetric equilibrium must be constrained efficient. Assume to the contrary that the ex ante expected utility in the symmetric unstable equilibrium was greater than in the asymmetric one by some $\delta > 0$. (Recall that we are considering the common interest case where income and satisfaction are public goods.) Let w_s denote the ex ante expected payoffs of men (which equals that of women) in the symmetric equilibrium. Consider now a small ε change in the men's threshold from the symmetric equilibrium. This causes an ex ante loss to everyone of order $\varepsilon < \delta$. Now allow women to best respond to this change of men's actions, and then men, and so on. This monotonic sequence converges to the stable asymmetric equilibrium. At each stage there is an ex ante gain for both so the payoff at the asymmetric equilibrium must be greater than $w_s - \varepsilon > w_s - \delta$, a contradiction.

The fact that when job satisfaction is a public good the equilibria are constrained Pareto efficient implies in turn that when job satisfaction is private the stable equilibrium is not ex ante constrained efficient and it is socially desirable to encourage individuals to choose the HIO.[19] As mentioned in section 4.2 this can result from various actions, such as decreasing the tax on households with two high-wage earners.

[18] This is because these men gain when averaging over the expected women with whom they are matched. Obviously women who are matched with other men have no change in their ex post utilities. Women matched with these men—whose actions have changed—do have a change in utility, and indeed some lose and some gain. But since the men gain when averaging over the women, and the payoffs are identical, the women – on average – gain as well.

[19] There may be other reasons to encourage choosing the HIO that come from sources outside the model. For example, one might believe that there are general externalities of productivity, such as tax revenues or spillovers.

6.4 Further extensions

There are several extensions of possible interest. One might initially be concerned that the results in this paper crucially rely on inefficiencies that bargaining or assortative matching may eliminate. While we think that analyzing these extensions is of interest, the arguments in sections 6.2 and 6.3 show that the externalities and inefficiencies do not exist when we allow job satisfaction to be public, while the equilibrium and its comparative statics are unchanged. Thus the inefficiencies are not inherent to our formal analysis and hence,we do not expect the results to change qualitatively by introducing bargaining or assortative matching just because they eliminate inefficiencies. Of course, bargaining and assortative matching will introduce other changes, including to our methodological results on stability, uniqueness and comparative statics. Similarly, adding a taste for discrimination explicitly, over and above our model, may interact with our results. Finally, we consider only two occupations; examining the extent to which our results extend to more occupations is of potential interest.

6.5 Alternative models

As we discussed we do not view the job-satisfaction gap as the whole story behind the wage gap; nevertheless we think it has an important role. Similarly, while the model above generates novel comparative statics, those do not require a job-satisfaction parameter—related models can achieve similar goals. In this subsection we discuss to what extent the main substantive results of the model—a job satisfaction gap combined with a wage gap, and comparative statics of the form discussed in Section 4.2—can be achieved in alternative models.

Obviously there can be many explanations for a job satisfaction gap. Consider a model without couples and where discrimination, in the from of a proportional reduction of salary for women, drives the wage gap. For this discrimination to be consistent with a job-satisfaction gap in favor of women it is necessary that the (long-run) substitution effect that pushes towards forgoing income for job satisfaction be stronger than the opposing income effect, in contrast to common belief. (In terms of the preferences the utility function cannot be too concave; for instance if the utility for money has constant relative risk aversion this coefficient must be less than 1.) By contrast in our model the wage and satisfaction gap combination relies on the role of couples and any degree of concavity is sufficient. Thus, accepting that the long-run income effect dominates, one way to distinguish our model from one of pure discrimination would be the existence of (the apparent) satisfaction gap.

Moreover, discrimination alone cannot yield the comparative statics we obtain, in particular those where a common change leads men and women to change behavior in opposite directions. We build these comparative statics on a model with couples, a tradeoff between income and job satisfaction and that income be jointly shared. However, our comparative statics results are more general: they do not require the tradeoff to be between income and job satisfaction; for example a tradeoff between higher income and flexible work hours that facilitate home production or with private costs of pre-marital investments (as in Peters and Siou (2002)) would lead to similar results. The comparative statics results are not intended to identify job satisfaction as the critical component, but rather to help identify implications of any one of a large class of similar models in which pre-match job choices drive wage gaps rather than discrimination alone.

7 References

Albanesi, Stefania and Claudia Olivetti (2009). "Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations", *Review of Economic Dynamics*, 12, 80–107.

Altonji, Joseph G. and Rebecca M. Blank (1999). "Race and Gender in the Labor Market", in Orley Ashenfelter and David Card eds, *Handbook of Labor Economics*, Volume 3, Part 3, Chapter 48, 3143–3259.

Bagnoli, Mark and Theodore Bergstrom (1993). "Courtship as a Waiting Game", *Journal of Political Economy*, 101, 185–202.

—— (2005). "Log-concave probability and its applications", Economic Theory, Springer, 26, (2) 445–469.

Becker, Gary S. (1993). "Division of Labor in Households and Families", in Gary Becker: *A Treatise on the Family*, Chapter 2, Cambridge, MA: Harvard University Press.

Bertrand, Marianne (2011). "New Perspectives on Gender", in Orley Ashenfelter and David Card eds, *Handbook of Labor Economics*, Volume 4b, Chapter 17.

Bhaskar V. and Ed Hopkins (2011), "Marriage as a Rat Race: Noisy Pre-Marital Investments with Assortative Matching," DP 8529 CEPR.

Booth, A. and M. Coles (2010). "Education, Matching and the Allocative Value of Romance," *Journal of the European Economic Association*, 8(4), 744–775.

Cole, Harold Linh, Mailath, George J., and Andrew Postlewaite (2001a). "Efficient Non-Contractible Investments in Large Economies", *Journal of Economic Theory*, 101, 333–373.

—— (2001b). "Efficient Non-Contractible Investments in Finite Economies", *Advances in Theoretical Economics*, 1, (1) Article 2.

—— (2001c). "Investment and Concern for Relative Position", *Review of Economic Design*, 6, (2) 241–261.

Danziger, Leif and Eliakim Katz (1996). "A theory of Sex Discrimination", *Journal of Economic Behavior & Organization* 31, (1) 57–66

Echevarria, Cristina and Antonio Merlo (1999). "Gender Differences in Education in a Dynamic Household Bargaining Model", *International Economic Review* 40 (2) 265–28.

Elul, Ronen, Jose Silva-Reus, and Oscar Volij (2002). "Will You Marry Me?: A Perspective on the Gender Gap", *Journal of Economic Behavior & Organization*, 49, (4) 549–572.

Engineer, Merwan, and Linda Welling (1998) "Human Capital, True Love, and Gender Roles: Is Sex Destiny?," *Journal of Economic Behavior and Organization*, **40** (2) 155–78

Felli, Leonardo and Kevin W.S. Roberts (2002). "Does Competition Solve the Hold-Up Problem?", CEPR Discussion Paper No. 3535.

Francois, Patrick (1998). "Gender Discrimination without Gender Difference: Theory and Policy Responses", *Journal of Public Economics*, 68, (1) 1–32.

Hadfield, Gillian K. (1999). "A Coordination Model of the Sexual Division of Labor", *Journal of Economic Behavior & Organization*, 40, 125–153.

Ishida, Junichiro (2003). "The Role of Social Norms in a Model of Marriage and Divorce", *Journal of Economic Behavior and Organization*, 51, (1) 131–142.

Iyigun, Murat and Randall P. Walsh (2007). "Building the Family Nest: Premarital Investments, Marriage Markets, and Spousal Allocations", *Review of Economic Studies*, 74, 507–535.

Lazear, Edward P. and Sherwin Rosen (1990). "Male-Female Wage Differentials in Job Ladders", *Journal of Labor Economics*, 8, (1) 106–23.

Lommerud, Kjell E. and Steinar Vagstad (2007). "Mommy Tracks and Public Policy: On

Self-fulfilling Prophecies and Mommy Tracks in Promotion". Mimeo.

Nosaka, Hiromi (2007). "Specialization and Competition in Marriage Models", *Journal of Economic Behavior & Organization*, 63, (1) 104–119.

Peters Michael and Aloysius Siow (2002). "Competing Premarital Investments", *Journal of Political Economy* 110, (3) 592–608.

Siow, Aloysius (1998). "Differential Fecundity, Markets, and Gender Roles", *Journal of Political Economy*, 106, (2) 334–354.

Sousa-Poza, Alfonso and Sousa-Poza, Andres A. (2003). "Gender Differences in Job Satisfaction in Great Britain, 1991-2000: Permanent or Transitory?", *Applied Economics Letters*, 10 (11) 691–694.

Stevenson, Betsey and Justin Wolfers (2009). "The Paradox of Declining Female Happiness", *American Economic Journal: Economic Policy*, 1, (2) 190–225.

Zafar, Basit (2008). "College Major Choice and the Gender Gap", Mimeo.

8 Appendix

Proposition 1 *Either there is a unique equilibrium x^s which is stable and symmetric with $|B'(x^s)| \leq 1$, or there are three equilibria: an unstable symmetric equilibrium x^s with $|B'(x^s)| > 1$ and two stable asymmetric equilibria (x, y) and (y, x) with $B'(x) \times B'(y) < 1$.*

Proof. Denote the best-reply function by $B(x) = q \cdot U^S + (1-q) \cdot \left[U^H F(x) + U^L (1 - F(x))\right]$. Clearly, $B(x) \in \left[\underline{b}, \overline{b}\right]$ where $\underline{b} = qU^S + (1-q) U^H > 0$ and $\overline{b} = qU^S + (1-q) U^L$ since if $F(x) = 1$ then anyone with dislike below $\underline{b}$ will choose the HIO, and if $F(x) = 0$ then anyone with dislike greater than $\overline{b}$ will choose the non-HIO. Since B is continuous, it has a fixed point in the closed interval $\left[\underline{b}, \overline{b}\right]$, which is a symmetric equilibrium. Consider now function $R(x) = B(B(x))$. Then in any equilibrium, symmetric or not, $x = R(x)$, i.e., equilibria are intersections of R with the 45-degree line. In a symmetric equilibrium $x = B(x) = R(x)$. An asymmetric equilibrium is a pair of thresholds (x, y) with $x = R(x)$ and $y = B(x) = R(y)$.

We consider R' at intersections $R(x) = x$, since $R' > 1$ implies instability of equilibrium (symmetric or not) and $R' < 1$ implies stability. (We will see below that $R' = 1$ implies stability.)

Since $U^H < U^L$ we have $B'(x) = (1-q)\left(U^H - U^L\right) f(x) < 0$. Since f is single peaked with peak below 0, then over the interval $[0, \bar{x}]$ we have $f'(x) < 0$ hence $B'' = (1-q)\left(U^H - U^L\right) f'(x) > 0$. That f is log-concave is equivalent to $\frac{f'(x)}{f(x)}$ being weakly decreasing, which implies that $\frac{B''(x)}{B'(x)}$ is weakly decreasing. This implies $0 > B''(x) B'(y) \geq B''(y) B'(x)$ for all $y > x$ (and $B''(x) B'(y) \leq B''(y) B'(x) < 0$ for all $y < x$).

Consider now a symmetric equilibrium $x^s = B(x^s)$. Since B is decreasing, for $x > x^s$ we have $x > x^s = B(x^s)$ and for $x < x^s$ we have $x < x^s = B(x^s)$. Furthermore,

$$R' = B'(B(x)) B'(x) = B'(y) B'(x) \tag{1}$$

$$R'' = B''(B(x)) (B'(x))^2 + B'(B(x)) B''(x) = B''(y) (B'(x))^2 + B'(y) B''(x). \tag{2}$$

Note that for an asymmetric equilibrium $R'(x) = R'(y)$. Also, since $\left[\underline{b}, \overline{b}\right] \subsetneq [0, \bar{x}]$ we have $f(x) > 0$ on $\left[\underline{b}, \overline{b}\right]$ so $B' < 0$ on $\left[\underline{b}, \overline{b}\right]$ and thus $R' > 0$ on $\left[\underline{b}, \overline{b}\right]$.

Consider the case where $R'(x^s) \geq 1$ and recall that

$$B''(y) (B'(x)) \leq B'(y) B''(x) \iff \tag{3}$$

$$|B''(y) (B'(x))| \geq |B'(y) B''(x)| \tag{4}$$

for $y > x > 0$. For $x < x^s$ (since $B'' > 0$ and $|B'(x^s)| = \sqrt{R'(x^s)} > 1$) we have $|B'(x)| > 1$. Hence multiplying the LHS of (3) by $B'(x)$ it becomes positive and is greater in absolute value than the RHS **and** hence $R''(x) > 0$. Thus,

$$R'(x^s) \geq 1 \Rightarrow R''(x) > 0 \ \forall x < x^s, \tag{5}$$

and similarly one can show

$$R'(x^s) \leq 1 \Rightarrow R''(x) < 0 \ \forall x > x^s. \tag{6}$$

First note that if $R'(x^s) = 1$ then x^s is stable. This is because when $R'(x^s) = 1$ we have from the preceding pair of equations that $R'(x) < 1$ for all $x \neq x^s$. So $B'(x) B'(B(x)) < 1$ which implies stability.

Thus x^s is stable iff $R'(x^s) \leq 1$ and then R does not cross the 45 degree line for any $x > x^s$ so there is no asymmetric equilibrium. (Recall that if there were an asymmetric equilibrium (x, y) then $R(x) = x$ and $R(y) = y$ and one of them would be greater than x^s and the other would be less.)

Also, x^s is unstable iff $R'(x^s) > 1$ and then R must cross the 45-degree line at some $\hat{x} < x^s$. (If not then for all $x < x^s$ we have $R(x) < x$ but this contradicts $R(\underline{b}) \geq \underline{b}$.) Thus $(\hat{x}, B(\hat{x}))$

is an asymmetric equilibrium. Moreover, since $R''(x) > 0$ for all $x < x^s$ this is the only x for which $R(x) = x$ and $R'(\hat{x}) < 1$ so it is the only asymmetric equilibrium with $x < x^s$ and it is stable. (Obviously there exists one other asymmetric equilibrium, its mirror image, $(B(\hat{x}), \hat{x})$.) ■

Lemma 2: In a stable equilibrium $(x^m(t), x^w(t))$, at $t = 0$,

$$sign(x_t^m(t)) = sign(B_t(x^w(t), t) + B_x(x^w(t), t) B_t(x^m(t), t))$$

and likewise

$$sign(x_t^w(t)) = sign(B_t(x^m(t), t) + B_x(x^m(t), t) B_t(x^w(t), t)).$$

Proof. Taking derivatives of $x^j = B(x^{-j}(t), t)$ wrt t we obtain:

$$\begin{aligned} x_t^m(t) &= B_t(x^w(t), t) + B_x(x^w(t), t) x_t^w(t) \\ x_t^w(t) &= B_t(x^m(t), t) + B_x(x^m(t), t) x_t^m(t) \end{aligned}$$

or in short, at $t = 0$ and dropping the variable t from $x^j(t)$

$$\begin{aligned} x_t^m &= B_t(x^w, 0) + B_x(x^w, 0) x_t^w \\ x_t^w &= B_t(x^m, 0) + B_x(x^m, 0) x_t^m \end{aligned}$$

and thus

$$x_t^m (1 - B_x(x^w, 0) B_x(x^m, 0)) = B_t(x^w, 0) + B_x(x^w, 0) B_t(x^m, 0).$$

By stability $1 - B_x(x^w, 0) B_x(x^m, 0) > 0$ so

$$sign(x_t^m) = sign(B_t(x^w, 0) + B_x(x^w, 0) B_t(x^m, 0))$$

and likewise

$$sign(x_t^w) = sign(B_t(x^m, 0) + B_x(x^m, 0) B_t(x^w, 0))$$

■

Proof of comparative-statics results in the table:

Recall that for $j = w, m$, the best-response function is:

$$x^j = B\left(x^{-j}\right) \equiv q \cdot U^S + (1 - q) \cdot \left[U^H F\left(x^{-j}\right) + U^L\left(1 - F\left(x^{-j}\right)\right)\right].$$

• U^S

The derivatives with respect to $t = U^S$ are $B_t^j = q > 0$, and thus $B_t^w(x^m) = B_t^m(x^w) > 0$. By Theorem 2 (2), the combined effects are $x_t^w > 0$ and $x_t^m < 0$.

- U^H

The derivatives with respect to $t = U^H$ are $B_t^j = (1-q) F(x^{-j}) > 0$. Since $F(x^m) > F(x^w)$ we thus have $B_t^w(x^m) > B_t^m(x^w) > 0$. By Theorem 2 (2), the combined effects are $x_t^w > 0$ and $x_t^m < 0$.

- U^L

The derivatives with respect to $t = U^L$ are $B_t^j = (1-q)(1 - F(x^{-j})) > 0$. Since $F(x^m) > F(x^w)$ we thus have $B_t^m(x^w) > B_t^w(x^m) > 0$. By Theorem 2 (3), at least one of the combined effects is positive, i.e., $x_t^w > 0$ or $x_t^m > 0$.

- q

The derivatives with respect to $t = q$ are $B_t^j = U^S - \left[U^H F(x^{-j}) + U^L(1 - F(x^{-j}))\right]$. If $U^S > \left[U^H F(x^m) + U^L(1 - F(x^m))\right]$ (which holds, in particular, if $U^S > U^H$), then $0 > B_t^w(x^m) > B_t^m(x^w)$ and by Theorem 2 (2), the combined effects are $x_t^w > 0$ and $x_t^m < 0$. If $U^S < \left[U^H F(x^m) + U^L(1 - F(x^m))\right]$ (which holds, in particular, if $U^S < U^L$), then $B_t^w(x^m) < B_t^m(x^w) < 0$ and by Theorem 2 (3), at least one of the combined effects is negative, i.e., $x_t^w < 0$ or $x_t^m < 0$. Note that if $\left[U^H F(x^m) + U^L(1 - F(x^m))\right] > U^S > \left[U^H F(x^m) + U^L(1 - F(x^m))\right]$, then $B_t^w(x^m) > 0 > B_t^m(x^w)$., and thus by Theorem 2 (1), the combined effects are $x_t^w > 0$ and $x_t^m < 0$.

- w_l

In order to take the derivatives with respect to $t = w_l$, note first that $U_{w_l}^H = -u_C'(w_h + w_l)$, $U_{w_l}^L = u_C'(w_h + w_l) - 2u_C'(w_l + w_l)$, and $U_{w_l}^S = -u_S'(w_l)$. We thus have:

$$B_t^w(x^m) = -q(u_S'(w_l)) + (1-q)\left[(-u_C'(w_h + w_l)) F(x^m) + (u_C'(w_h + w_l) - 2u_C'(w_l + w_l))(1 - F(x^m))\right] < 0$$

$$B_t^m(x^w) = -q(u_S'(w_l)) + (1-q)\left[(-u_C'(w_h + w_l)) F(x^w) + (u_C'(w_h + w_l) - 2u_C'(w_l + w_l))(1 - F(x^w))\right] < 0.$$

(Here both inequalities hold because $u_C'(w_l + w_l) > u_C'(w_h + w_l)$.) Thus,

$$|B_t^w(x^m)| - |B_t^m(x^w)| = B_t^m(x^w) - B_t^w(x^m)$$
$$= (1-q)\left[(-u_C'(w_h + w_l))(F(x^w) - F(x^m)) + (u_C'(w_h + w_l) - 2u_C'(w_l + w_l))((1 - F(x^w)) - (1 - F(x^m)))\right]$$
$$= (1-q)\left[(-u_C'(w_h + w_l))(F(x^w) - F(x^m)) - (u_C'(w_h + w_l) - 2u_C'(w_l + w_l))(F(x^w) - F(x^m))\right]$$
$$= (1-q)\left[(F(x^w) - F(x^m))\, 2\left((-u_C'(w_h + w_l)) + u_C'(w_l + w_l)\right)\right] < 0$$

(again because $u'_C(w_l + w_l) > u'_C(w_h + w_l)$.) Therefore the direct effect on males is larger (more negative), and by Theorem 2 (3), at least one of the combined effects is negative, i.e., $x_t^w < 0$ or $x_t^m < 0$.

- w_h

To take the derivatives with respect to $t = w_h$, note first that $U_{w_h}^H = 2u'_C(w_h + w_h) - u'_C(w_h + w_l)$, $U_{w_h}^L = u'_C(w_h + w_l)$, and $U_{w_h}^S = u'_S(w_h)$. We thus have:
$B_t^j = q\left(u'_S(w_h)\right) + (1-q)\left[\left(2u'_C(w_h + w_h) - u'_C(w_h + w_l)\right)F\left(x^{-j}\right) + \left(u'_C(w_h + w_l)\right)\left(1 - F\left(x^{-j}\right)\right)\right]$.

Without further assumptions, this cannot be signed: because $u'_C(w_h + w_h) < u'_C(w_h + w_l)$ it can be that $2u'_C(w_h + w_h) - u'_C(w_h + w_l) < 0$ and then the sign of the overall expression can only be determined when the values of the parameters are known. However, if $w_l \geq 0$ and the coefficient of relative risk aversion is less than 1, then the sign of $2u'_C(w_h + w_h) - u'_C(w_h + w_l)$ is positive and hence $B_t^j > 0$. However in this case $B_t^w(x^m) < B_t^m(x^w)$ so all we can say is that the overall effect is only weakly determined (Theorem 2 (3)).

- Post-divorce share α

Modify the model so that a divorcee with wage w_j, whose partner had wage w_{-j}, will have income $(1-\alpha)w_j + \alpha w_{-j}$. Thus we define the post-divorce incentives for taking the HIO:

$$\begin{aligned} U^{D,L} &= u_S\left((1-\alpha)w_h + \alpha w_l\right) - u_S\left((1-\alpha)w_l + \alpha w_l\right) \\ U^{D,H} &= u_S\left((1-\alpha)w_h + \alpha w_h\right) - u_S\left((1-\alpha)w_l + \alpha w_h\right). \end{aligned}$$

The best response function is, in thus case,

$$x^j = q\left(U^{D,L}\left(1 - F\left(x^{-j}\right)\right) + U^{D,H}F\left(x^{-j}\right)\right) + (1-q)\cdot\left[U^H F\left(x^{-j}\right) + U^L\left(1 - F\left(x^{-j}\right)\right)\right].$$

Hence:

$$
\begin{aligned}
\frac{dx^j}{d\alpha} &= \left(u_S'\left(\left(1-\alpha\right)w_h+\alpha w_l\right)\left(w_l-w_h\right)\right)\left(1-F\left(x^{-j}\right)\right)+ \\
&\quad \left(-u_S'\left(\left(1-\alpha\right)w_l+\alpha w_h\right)\right)\left(w_h-w_l\right)F\left(x^{-j}\right) \\
&= -\left(u_S'\left(\left(1-\alpha\right)w_h+\alpha w_l\right)\right)\left(w_h-w_l\right)\left(1-F\left(x^{-j}\right)\right)+ \\
&\quad \left(-u_S'\left(\left(1-\alpha\right)w_l+\alpha w_h\right)\right)\left(w_h-w_l\right)F\left(x^{-j}\right) \\
&= -\left(w_h-w_l\right)\left[\begin{array}{c} \left(u_S'\left(\left(1-\alpha\right)w_h+\alpha w_l\right)\right)\left(1-F\left(x^{-j}\right)\right)+ \\ \left(u_S'\left(\left(1-\alpha\right)w_l+\alpha w_h\right)\right)F\left(x^{-j}\right) \end{array}\right] \\
&< 0
\end{aligned}
$$

Since $B_\alpha^j\left(x^{-j}\right) < 0$ we compare $\left|B_\alpha^w\left(x^m\right)\right|$ with $\left|B_\alpha^m\left(x^w\right)\right|$. Note that $\alpha \in [0,1/2]$ so $u_S'\left(\left(1-\alpha\right)w_h+\alpha w_l\right) < u_S'\left(\left(1-\alpha\right)w_l+\alpha w_h\right)$. As $F\left(x^m\right) > F\left(x^w\right)$ we have

$$
\begin{aligned}
\left|\frac{dx^m}{d\alpha}\right| &= \left(w_h-w_l\right)\left(\left(u_S'\left(\left(1-\alpha\right)w_h+\alpha w_l\right)\right)\left(1-F\left(x^w\right)\right)+\left(u_S'\left(\left(1-\alpha\right)w_l+\alpha w_h\right)\right)F\left(x^w\right)\right) \\
&< \left(w_h-w_l\right)\left(\left(u_S'\left(\left(1-\alpha\right)w_h+\alpha w_l\right)\right)\left(1-F\left(x^m\right)\right)+\left(u_S'\left(\left(1-\alpha\right)w_l+\alpha w_h\right)\right)F\left(x^m\right)\right) \\
&= \left|\frac{dx^w}{d\alpha}\right|
\end{aligned}
$$

Therefore as post divorce income is more equally divided, women choose HIO less often and men choose it *more* often (Theorem 2 (2)).