Taxation of Couples under Assortative Mating

By Alexander Frankel

I present a simple and tractable model of the optimal taxation of married couples, working off of the multidimensional screening framework of Armstrong and Rochet (1999). In particular, I study how the tax code varies with the degree of assortative mating. One result is that the “negative jointness” of marginal tax rates found in Kleven, Kreiner, and Saez (2007, 2009) for couples with uncorrelated earnings should be attenuated in the presence of assortative mating. When mating is sufficiently assortative, the optimal tax schedule is separable: an individual’s taxes do not depend on his or her spouse’s income. (JEL D82, H21, H24, J12)

As well-off men become increasingly likely to marry well-off women, inequality across couples goes up. This relationship between assortative mating and income inequality has been discussed by Burtless (1999), Cancian and Reed (1999), Schwartz (2010), and others. Moreover, as documented by Schwartz and Mare (2005), assortative mating is prominent in the United States and has been increasing in recent decades. Surprisingly enough, very little attention has been paid to the interaction of assortative mating with the natural policy response to inequality: redistributive taxation and transfers.

I consider a simple model of optimal couples taxation in order to study this interaction. A population of couples is determined; the government sets a tax and transfer schedule; individuals then choose how much to earn. Taxes may be an arbitrary function of the output of each individual in the couple. The key simplifying assumption of the model is that men and women in a couple each have a binary productivity type, either high or low. So mating becomes more assortative as we reduce the number of mixed high-low and low-high couples, and replace them with matched high-high and low-low couples. The government has a preference for redistributing money to the worst-off couples, the ones with men and women who are both of low types.

This stripped-down model of taxation is quite tractable. I show that the government’s problem can be transformed into the monopoly regulation problem of Armstrong and Rochet (1999). The results of Armstrong and Rochet (1999) imply output levels of each individual under the optimal tax schedule. These output levels

* University of Chicago Booth School of Business, 5807 S Woodlawn Ave, Chicago IL 60637 (e-mail: afrankel@uchicago.edu). I would like to thank Jeremy Bulow, Caroline Hoxby, Michael Lovenheim, and Robert Wilson for their guidance on this project, as well as Kory Kroft, Matthew Notowidigdo, and Alessandra Voena for useful advice. The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

† Go to http://dx.doi.org/10.1257/pol.6.3.155 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.
can then be used to characterize the corresponding tax schedule, and to study how it varies with the level of assortative mating. The solution does not require any functional form assumptions on individuals’ costs of effort, and allows for asymmetries across genders.

One of the main contributions of the paper is to shed light on the issue of “jointness” in the tax code. When my spouse earns more money, should my marginal tax rate increase, decrease, or stay the same? Kleven, Kreiner, and Saez (2007, 2009) present models of couples taxation which argue for negative jointness: reducing one’s marginal tax rate when his or her partner’s income goes up. They prove the optimality of negative jointness in models with continuously distributed types, compared to the binary types of the current paper, but in which couples are matched independently: low and high type men are equally likely to marry a high type woman. My simpler binary model allows for the analysis of assortative mating, i.e., couples with correlated types. I reproduce the negative jointness result when one’s type is uncorrelated to his or her spouse’s. As we increase the level of assortative mating, though, negative jointness should be attenuated. For a large enough correlation between types, I find that negative jointness disappears entirely. In an economy where couples have highly correlated types, the government should use a separable tax code, one where an individual’s taxes do not depend at all on his or her spouse’s income.

The design of optimal redistributive tax schedules for individuals, modeled as a one-dimensional screening problem, has been studied extensively since Mirrlees (1971). There has been much less work on the problem of optimal couples taxation as a two-dimensional screening problem, in which both individuals have a distinct productivity type, and the government can set arbitrary taxes as a function of the two spouses’ outputs. This literature includes the Kleven, Kreiner, and Saez (2007, 2009) papers mentioned above; I discuss these papers further in Section IE. Another recent paper in this space is Brett (2007), which proposes a model of couples taxation similar to the current one: each individual has a binary productivity type, and couples act as unitary decisionmakers. Indeed, his mathematical formulation is more general than my own, in that it allows for a broader set of social welfare functions and allows for couples with a concave utility over income; couples in my model have quasilinear utility in income. However, this general formulation prevents him from solving for the optimal tax code. Instead, he derives restrictions on the sets of incentive compatibility constraints which may bind at an optimum, and discusses how a given set of binding constraints translates into marginal tax rates and output levels for each type of individual in each type of couple.

In other related work, Immervoll et al. (2011) consider how to optimally tax or subsidize labor force participation when each couple has a one-dimensional continuous type determining the productivity of both the primary and secondary earner. Cremer, Lozachmeur, and Pestieau (2012) consider a model of couples taxation and allow for general discrete type distributions. They look for conditions under which the two spouses should face identical tax rates, i.e., for when the government taxes only the sum of a couple’s income. Immervoll et al. (2009) conduct an empirical study of the jointness of tax codes across countries and quantify the extent to which there is a “marriage penalty.”
I. The Model

There is a population of couples which earn income, and a government which taxes and redistributes this income.

A. The Couple’s Problem

A couple is made up of a pair of individuals, which I call a man $m$ and a woman $w$. Individual $i = m, w$ in a couple has an earnings type $\theta^i$, and chooses output (pre-tax earnings) of $Y^i \geq 0$ at an effort cost of $C^i(Y^i|\theta^i)$. The couple then pays a joint tax $\tau(Y^m, Y^w)$. This tax payment can depend arbitrarily on both individuals’ pretax incomes, and it may be negative, indicating a transfer. Given a tax schedule $\tau(\cdot, \cdot)$ along with types $\theta^m$ and $\theta^w$, the couple’s problem is to choose $Y^m$ and $Y^w$ to maximize utility equal to after-tax earnings net of effort costs:

$$U = \sum_{i=m,w} (Y^i - C^i(Y^i|\theta^i)) - \tau(Y^m, Y^w).$$

Couples here act as unitary decision makers, choosing each individual’s output to maximize joint utility. Couples also have quasilinear utility in money. In other words, there are no income effects in labor participation. If a husband earns more, it only affects the wife’s earnings decision by potentially changing the marginal tax rate on her output.

I assume that individual types are binary: each $\theta^i$ is either equal to $L$ or $H$. The types $L$ and $H$ are “low” and “high” in the sense that the high types have lower marginal costs of earning money. High types therefore choose to earn more than low types, all else equal. Formally, writing marginal effort costs as $m_c^i(y|\theta^i) \equiv \frac{\partial}{\partial y} C^i(y|\theta^i)$, I assume that $m_c^i(y|H) < m_c^i(y|L)$ for each $i = m, w$ and each $y > 0$.

As a matter of notation, I say that the partner of individual $i$ has type $\theta^{-i}$. The couple’s joint type is $\Theta = \theta^m \theta^w$, with $\Theta$ an element of the joint type space $\Theta = \{HH, HL, LH, LL\}$. I denote a couple of type $\Theta$’s outputs, tax, and utility values as $Y^i_{\theta^i}$, $\tau_{\theta}$, and $U_{\theta}$.

Throughout the paper I impose the following regularity conditions on the cost functions: for each $i$, $MC^i(y|\theta^i)$ is well-defined, continuous, and strictly increasing in $y$; $0 \leq MC^i(0|\theta^i) \leq 1$; and $MC^i(y|\theta^i) > 1$ for $y$ large enough. Let $y^i_{\theta^i}$ be the point such that $MC^i(y^i_{\theta^i}|\theta^i) = 1$, the efficient level of individual output. By the regularity conditions, $y^i_{\theta^i}$ is positive and uniquely defined.

B. The Population of Couples

In the population, there is a mass $\alpha_{\theta} \geq 0$ of couples of joint type $\theta \in \Theta$. The total mass of couples is normalized to one: $\sum_{\theta} \alpha_{\theta} = 1$. The distribution of couples is determined exogenously and does not depend on the tax code.
Given a distribution $\alpha$, I measure the level of assortative mating in the economy using the parameter $\rho$, which is akin to a covariance:

$$\rho \equiv \alpha_{LL}\alpha_{HH} - \alpha_{LH}\alpha_{HL}.$$ 

A value of $\rho = 0$ indicates that the man’s type is statistically uncorrelated with the woman’s type, while $\rho > 0$ indicates a positive correlation. Holding the number of high and low individuals of each gender fixed, the entire joint distribution of types is uniquely determined by $\rho$. In particular, say that a proportion $q_i$ of gender $i$ are high types and $1 - q_i$ are low. Then $\alpha_{LL} = 1 - q^m_i q^w_i + q^m_i q^w_i \rho$; $\alpha_{HH} = q^m_i q^w_i + \rho$; $\alpha_{HL} = q^m_i - q^m_i q^w_i - \rho$; and $\alpha_{LH} = q^w_i - q^m_i q^w_i - \rho$. As we increase $\rho$, there are more matched low-low and high-high couples, and fewer mixed couples.

This paper will focus on the case of weakly assortative mating, $\rho \geq 0$. As I describe in later sections, it will be useful to distinguish between two regions of correlation. For $0 \leq \rho \leq \alpha_{LH}\alpha_{HL}/\alpha_{LL}$, I say that there is Weak Positive Correlation. For $\rho > \alpha_{LH}\alpha_{HL}/\alpha_{LL}$, there is Strong Positive Correlation. Walking through these definitions, suppose we start in a distribution with $\rho = 0$. It must hold that $\alpha_{LH}\alpha_{HL}/\alpha_{LL} \geq 0$, so we satisfy Weak Positive Correlation. If we then reduce the number of $LH$ and $HL$ couples and replace them with $HH$ and $LL$ couples, this increases $\rho$ and at the same time decreases $\alpha_{LH}\alpha_{HL}/\alpha_{LL}$. Continuing in this manner, eventually $\rho \geq \alpha_{LH}\alpha_{HL}/\alpha_{LL}$ and we switch to the Strong Positive Correlation regime.

In the special case where both genders are equally split between high and low types, it holds that $\alpha_{HH} = \alpha_{LL} = 1/4 + \rho$ and $\alpha_{HL} = \alpha_{LH} = 1/4 - \rho$, for $\rho \in [-1/4, 1/4]$. The statistical correlation coefficient between types in the population is exactly $4\rho$. Under this even split of high and low types, we switch from Weak to Strong Positive Correlation at $\rho = 1/12$, or a correlation coefficient of $1/3$.

While most of the analysis of this paper will allow for general assortative type distributions $\alpha$ and general asymmetric cost functions $C^i$, at times it will be convenient to make the simplifying assumption that genders are fully symmetric. Letting $q \in (0, 1)$ represent the proportion of high types in each gender, a symmetric type distribution is equivalent to $\alpha_{HL} = \alpha_{LH} = q - q^2 - \rho$, $\alpha_{HH} = q^2 + \rho$, and $\alpha_{LL} = 1 - 2q + q^2 + \rho$.

**ASSUMPTION 1 (Symmetric Genders):** Let cost functions satisfy $C^m(y) = C^w(y)$ for all $y$, and let the distribution of types satisfy $\alpha_{HL} = \alpha_{LH} = q - q^2 - \rho$, $\alpha_{HH} = q^2 + \rho$, and $\alpha_{LL} = 1 - 2q + q^2 + \rho$ for some $q \in (0, 1)$.

Under Symmetric Genders, we have Weak Positive Correlation for $0 \leq \rho \leq q^2(1 - q)/(1 + q)$ and Strong Positive Correlation for $q^2(1 - q)/(1 + q) < \rho \leq q - q^2$. The cutoff of $\rho = q^2(1 - q)/(1 + q)$ corresponds to a correlation coefficient of joint types of $q/(1 + q)$. I will be explicit in this paper about when the assumption of Symmetric Genders is and is not imposed.

---

1 The notation here and in many other places is borrowed from Armstrong and Rochet (1999).
2 Any value of $\rho$ is feasible so long as each induced $\alpha_{H}$ value is nonnegative.
C. The Government’s Problem

The government’s problem is to use redistributive taxation to maximize the weighted average population utility, with the LL types taking a weight of 1 while all other types take a weight of $\lambda \in [0, 1]$. We can write the social welfare function as

$$SW = \frac{\alpha_{LL} U_{LL} + \lambda \sum_{\theta \neq LL} \alpha_{\theta} U_{\theta}}{\alpha_{LL} + \lambda (1 - \alpha_{LL})}. $$

The parameter $\lambda$ controls the government’s preference for redistribution. A larger $\lambda$ corresponds to a smaller desire for redistribution. Setting $\lambda = 1$ means that the government maximizes average population utility, while $\lambda = 0$ is Rawlsian, maximizing the welfare of the worst-off types.\(^3\)

The government chooses a tax schedule $\tau(Y^m, Y^w)$ as a function of the man’s and woman’s output in a couple. (It can observe genders and outputs of each individual, but cannot observe types.) Couples take this tax policy as given when choosing output levels; the government is the first mover. The government must respect a balanced budget condition, in that the sum of all couples’ tax payments must be weakly positive.

The government’s problem is equivalent to choosing a pair of outputs $(Y^m_{\theta}, Y^w_{\theta})$ and a tax $\tau_{\theta}$ for each joint type $\theta \in \Theta$ subject to incentive compatibility (IC) and balanced budget (BB) constraints. The IC constraint that $\theta$ does not pretend to be type $\theta'$ is

$$ (IC_{\theta, \theta'}) \sum_{i=m,w} (Y^i_{\theta} - C^i(Y^i_{\theta} | \theta^i)) - \tau_{\theta} \geq \sum_{i=m,w} (Y^i_{\theta'} - C^i(Y^i_{\theta'} | \theta^i)) - \tau_{\theta'}. $$

At an optimal tax schedule, the balanced budget constraint will bind. Otherwise the government could reduce taxes for all couples without affecting incentive compatibility. So the BB constraint can be written as

$$ (BB) \sum_{\theta \in \Theta} \alpha_{\theta} \tau_{\theta} = 0. $$

I sometimes use the notation $E_{\alpha}$ to indicate expectation over types, given the distribution $\alpha$. For instance, the budget balance constraint is $E_{\alpha}[\tau_{\theta}] = 0$.

Substituting the expression for $U_{\theta}$ and $U_{\theta'}$ into the IC constraints, the government’s problem is

$$ \max_{\{Y^m_{\theta}, \tau_{\theta}, U_{\theta}\}} \quad \frac{\alpha_{LL} U_{LL} + \lambda \sum_{\theta \neq LL} \alpha_{\theta} U_{\theta}}{\alpha_{LL} + \lambda (1 - \alpha_{LL})} $$

s.t. $U_{\theta} = \sum_{i=m,w} (Y^i_{\theta} - C^i(Y^i_{\theta} | \theta^i)) - \tau_{\theta}$

\(^3\)Under any tax code, low-low couples will have weakly lower utility than other couples.
\[ (IC_{\theta, \theta'}) \quad U_\theta \geq U_{\theta'} + \sum_{i=m,w} \left( C^i(Y^i_{\theta'} | \theta'^i) - C^i(Y^i_{\theta} | \theta^i) \right) \quad \forall \theta, \theta' \]

\[ (BB) \quad E_\alpha[\tau_\theta] = 0. \]

D. Properties of a Tax Code

Marginal tax rates are not given directly by this model, since an optimal tax schedule will only pin down tax levels at four discrete joint output levels. We can determine imputed marginal tax rates from the output levels, however. If the couple faced a smooth tax schedule \( \tau(Y^m, Y^w) \) and was free to choose output levels, it would choose \( Y^i \) to satisfy the first order condition \( MC^i(Y^i | \theta^i) = 1 - \frac{\partial \tau(Y^m, Y^w)}{\partial Y^i} \).

This gives an imputed marginal tax rate of \( \frac{\partial \tau(Y^m, Y^w)}{\partial Y^i} = 1 - MC^i(Y^i_{\theta} | \theta^i) \) for individual \( i \) who chooses output \( Y^i_{\theta} \) from the menu provided by the government. I denote this imputed tax rate by \( t^i_{\theta} \):

\[ t^i_{\theta} \equiv 1 - MC^i(Y^i_{\theta} | \theta^i). \]

Following Kleven, Kreiner, and Saez (2007, 2009), one key focus will be on the “jointness” of this tax code. This describes how an individual’s marginal tax rates increase or decrease in the earnings of his or her partner. Positive jointness means that tax rates go up in partner’s earnings, and negative jointness means that tax rates decline in partner’s earnings. Conditional on one’s own type, one’s partner’s earnings are monotonic in his or her type. So it is equivalent and more convenient to here define jointness as the change in marginal taxes with respect to his or her partner’s type, rather than his or her partner’s earnings.

DEFINITION: Say that a tax system has zero jointness if \( t^i_{\theta} = t^i_{\theta'} \) whenever \( \theta^i = \theta'^i \). It has negative jointness if \( \theta^i = \theta'^i, \theta^i - \theta'^i = L, \) and \( \theta'^i - \theta^i = H \) imply that \( t^i_{\theta} \geq t^i_{\theta'} \). It has positive jointness if this last inequality is reversed.

Another property to consider about a tax code is whether it is separable, i.e., whether individuals are taxed independently of their partner’s income. A continuous tax schedule \( \tau(Y^m, Y^w) \) would be separable if there existed functions \( \tau^m(Y^m) \) and \( \tau^w(Y^w) \) such that \( \tau(Y^m, Y^w) = \tau^m(Y^m) + \tau^w(Y^w) \) for all \( Y^m, Y^w \). In this discrete model, a tax code is separable if the imputed marginal tax rates \( t^i_{\theta} \) and tax payments \( \tau_\theta \) are consistent with some continuous separable schedule.

DEFINITION: A tax schedule is separable if it has zero jointness, and if there exist \( \tau^i_{\theta^i} \) for each \( i = m, w \) and each \( \theta^i = L, H \) such that \( \tau^i_{\theta^i} = \tau^i_{\theta^m} + \tau^i_{\theta^w} \).

When the tax code is separable, a couple’s net tax bill \( \tau_\theta \) can be allocated across the two individuals as \( \tau^m_{\theta^m} + \tau^w_{\theta^w} \) in a number of ways. Reducing each man’s tax level by a dollar while increasing each woman’s tax level by a dollar, for instance, does not affect any couple’s tax level. One way to pin down the individual taxes
τ_θ^i is to suppose that taxes are budget balanced within each gender: no net subsidies from men to women, or women to men. Given any separable tax schedule, there is a unique collection of individual taxes τ_θ^i which is budget balanced across genders.

Note that zero jointness by itself does not imply separability. This is because the jointness pins down only the marginal tax rates, but not the tax levels. For instance, a zero joint tax code could be consistent with tax levels τ_{HH} = 10, τ_{HL} = τ_{LH} = 0, and τ_{LL} = −8. Comparing τ_{HH} and τ_{HL}, we see that a woman married to a high type man pays $10 more in taxes if her type is high rather than low. Comparing τ_{LH} and τ_{LL}, though, when married to a low type man this difference is $8. The two differences would have to be identical under any separable tax schedule, both equal to τ^w_{H} − τ^w_{L}. Indeed, separability is equivalent to zero jointness combined with τ_{HH} − τ_{HL} = τ_{LH} − τ_{LL} and τ_{HH} − τ_{LH} = τ_{HL} − τ_{LL}.

E. Relationship to Kleven, Kreiner, and Saez (2007, 2009)

The model of couples taxation in this paper is similar to the ones considered in Kleven, Kreiner, and Saez (2007, 2009). The key difference between my model and those of Kleven, Kreiner, and Saez (2007, 2009) is that I assume that individual types are binary (high or low), while Kleven, Kreiner, and Saez (2007, 2009) take them to be continuous.

Because a couple’s utility is quasilinear in income in all of these models, a social planner maximizing the sum of total welfare would simply eliminate all taxes. To give the government a preference for redistribution, Kleven, Kreiner, and Saez (2007, 2009) have the government maximize a concave transformation of average couple utility: \(E[Ψ(U_θ)]\), for some concave function Ψ. The current paper achieves redistribution by supposing that the government maximizes weighted average population utility where couples of lower type (who, in equilibrium, will have lower utility) are given a higher Pareto weight. In particular, the government places a weight of 1 on low-low couples and a weight of \(λ \leq 1\) on all other couples. This is equivalent in the Kleven, Kreiner, and Saez (2007, 2009) framework of choosing a piecewise linear Ψ, with slope 1 below some cutoff and slope λ above.

Kleven, Kreiner, and Saez (2007) present two distinct models of couples taxation. The first, which is repeated in the published version of Kleven, Kreiner, and Saez (2009), is an intensive-extensive model. The primary earner and secondary earner are fundamentally asymmetric: the primary earner makes a continuous effort choice, while the secondary earner chooses only whether or not to work. The second is a doubly intensive model, where the primary and secondary earner both make continuous effort choices. My paper can be thought of as a two-by-two version of the doubly intensive model.

In both models, Kleven, Kreiner, and Saez show that the optimal tax code features negative jointness if individual types are independently distributed[^4]. One should face a lower tax rate if his or her spouse earns more. For more general joint type

[^4]: See Kleven, Kreiner, and Saez (2009, proposition 3, part 1b) for the intensive-extensive result, and Kleven, Kreiner, and Saez (2007, theorem 1) for the doubly intensive result. For these results the authors require that the derivative of their social welfare function, Ψ(τ), be convex; see footnote 6, below.
distributions, they provide first-order conditions for optimal tax schedules, but do not have analytic results on jointness.

The authors give the following variational logic for negative jointness under uncorrelated mating. Start with a separable tax code, and consider the following perturbation towards negative jointness. Pick some earnings threshold for men and some earnings threshold for women. Slightly reduce the level of taxes paid by couples with two high earners (above the threshold) and with two low earners (below the threshold), while increasing the taxes on couples with a high and a low earner. Equivalently, we make a redistributive transfer from low-high to low-low couples while countering this with an anti-redistributive transfer from high-low to high-high couples. This introduces a small amount of negative jointness into the tax code; a woman near the threshold faces a lowered marginal tax rate if her husband is a high earner, and an increased one if her husband is low.

We can break the welfare impact of the perturbation into two parts: the direct welfare effect of the transfers, taking earnings choices as given; and the indirect fiscal effect due to the earnings responses of men and women. Consider first the direct effect. When the social planner has sufficiently redistributive preferences, the benefit of a transfer from low-high to low-low couples is greater than the harm of an equal transfer from high-low to high-high couples. Then Kleven, Kreiner, and Saez (2007, 2009) show that there is no first order fiscal cost or benefit, assuming independent types. So the perturbation towards negative jointness increases social welfare.

This analysis raises the question of whether the optimality of negative jointness is robust to assortative mating. Unfortunately, their arguments do not give intuition for how the result might change as we add correlation of the husband’s and wife’s earnings types. Kleven, Kreiner, and Saez (2007, 2009) address this issue in some limited numerical simulations of the intensive-extensive model. They add a form of assortative mating by allowing the support of the secondary earner’s type distribution to vary with the primary earner’s type. They find that adding correlation in this manner does not alter the qualitative features of the optimal negatively joint tax code.

In my binary framework, I can study correlation straightforwardly: we replace HL and LH couples with LL and HH couples, keeping the total proportion of high type men and high type women in the population constant. I will confirm that my model replicates the negative jointness of the tax code under no correlation. However, negative jointness is attenuated as correlation increases. For high enough correlations, there is a separable tax code: one’s taxes are independent of his or her spouse’s earnings. (See Propositions 2, 3, and 5.) In Section IIIIB, I walk through a variational

---

5The interested reader should refer to the original papers for the details of the argument, or to an earlier draft (Kleven, Kreiner, and Saez 2006).

6Kleven, Kreiner, and Saez (2007, 2009) guarantee that preferences are sufficiently redistributive by assuming that Ψ is convex. (With concave Ψ, the tax code becomes positively joint.) In my analogous two-by-two model, we can think of the welfare weight of 1 on low-low couples and λ ≤ 1 on other couples as implied values of Ψ. A convex Ψ corresponds to a higher value of high-high minus mixed welfare weights (λ − λ = 0) than mixed minus low-low (λ − 1 < 0). A high-low to high-high transfer is welfare neutral while a low-high to low-low transfer is positive.

7Somewhat surprisingly, one’s partner’s activity is ignored by the tax system exactly when it is most informative about one’s own type. Armstrong and Rochet (1999) have previously pointed out this counterintuitive feature of multidimensional screening in the context of monopoly price regulation; see their footnote 6.
argument similar to the ones of Kleven, Kreiner, and Saez (2007, 2009) to give intuition for the switch from negative to zero jointness under assortative mating.

II. Output Levels in the Optimal Tax Schedule

The problem of solving for the optimal tax code involves multidimensional screening: couples have two-dimensional private information on the individual types, and take a two-dimensional action. In this section I will show how to transform the variables of my model to put the government’s problem into the mathematical framework of the multidimensional monopoly pricing model of Armstrong and Rochet (1999). I can then apply results from that paper to derive output levels and marginal tax rates.

Let us rewrite social welfare as

$$SW = \frac{\alpha_{LL} U_{LL} + \lambda \sum_{\theta \neq LL} \alpha_{\theta} U_{\theta}}{\alpha_{LL} + \lambda (1 - \alpha_{LL})}$$

$$= \frac{\alpha_{LL} U_{LL} + \lambda (1 - \alpha_{LL}) U_{LL} + \lambda \sum_{\theta \in \Theta} \alpha_{\theta} (U_{\theta} - U_{LL})}{\alpha_{LL} + \lambda (1 - \alpha_{LL})}$$

$$= U_{LL} + \frac{\lambda}{\alpha_{LL} + \lambda (1 - \alpha_{LL})} E_{\alpha}[U_{\theta} - U_{LL}].$$

Now replace $U$, $\tau$, and $\lambda$ with the new variables, $R$, $T$, and $\beta$. Let $R_{\theta} \equiv U_{\theta} - U_{LL}$, $T_{\theta} \equiv \tau_{\theta} + U_{LL}$, and $\beta \equiv \lambda/(\alpha_{LL} + \lambda (1 - \alpha_{LL}))$. The budget balance condition (BB) is equivalent to supposing that $E_{\alpha}[T_{\theta}] = U_{LL}$. Under the budget balance assumption, then, $SW = E_{\alpha}[T_{\theta} + \beta R_{\theta}]$. In terms of the new variables, the government’s problem is

$$\max_{\{y_{\theta}, T_{\theta}, R_{\theta}\}} E_{\alpha}[T_{\theta} + \beta R_{\theta}]$$

$$\text{s.t.} \quad R_{LL} = 0$$

$$R_{\theta} = \sum_{i=m,w} (Y_{\theta} - C^{i}(Y_{\theta}^{i} | \theta^{i})) - T_{\theta}$$

$$(IC_{\theta, \theta}) \quad R_{\theta} \geq R_{\theta'} + \sum_{i=m,w} (C^{i}(Y_{\theta'}^{i} | \theta'^{i}) - C^{i}(Y_{\theta'}^{i} | \theta^{i})) \quad \forall \theta, \theta'.$$

We have removed the (BB) condition from the program, since (BB) is implicit in the expression for Social Welfare. We have also added the condition $R_{LL} = 0$, which holds from the previous definition $R_{\theta} \equiv U_{\theta} - U_{LL}$. 
This program is exactly the one considered in Armstrong and Rochet (1999) in a context of a monopolist selling two-dimensional goods to consumers, where in that paper $T_\theta$ represents net payments and $R_\theta$ the resulting utility of consumers with preference type $\Theta$. The monopolist places a weight of $\beta$ on consumer welfare. The expression $R_{LL} = 0$ represents the binding individual rationality constraint on low types to be willing to enter the market; for all other types, individual rationality is implied by incentive compatibility.8

Armstrong and Rochet (1999) characterize the general solution for the optimal output levels (in their model, quantities purchased) $Y^i_\theta$. To express these optimal output levels, define $\phi^i(\zeta)$ and $\hat{y}^i(\zeta)$ for $\zeta \geq 0$ in the following manner:

$$\phi^i(\zeta) \equiv \max_{y \geq 0} y - C^i(y | L) - \zeta \left(C^i(y | L) - C^i(y | H)\right),$$

and take $\hat{y}^i(\zeta)$ to be the arg max of the above expression. The undistorted earnings choice for a low type, $y^*_L$, corresponds to $\hat{y}^i(0)$. The function $\hat{y}^i(\zeta)$ is continuous and weakly decreasing in $\zeta$, and the function $\phi^i(\zeta)$ is weakly decreasing and convex in $\zeta$.

**THEOREM 1** (From Armstrong and Rochet 1999):

**Weak Positive Correlation.** If $0 \leq \rho \leq \frac{\alpha_{HH}^* \alpha_{HL}}{\alpha_{LL}}$, then the optimal policy induces output levels

\begin{align*}
Y^m_{HH} &= Y^m_{HL} = y^m_L, & Y^w_{HH} &= Y^w_{HL} = y^w_H \\
Y^m_{LL} &= \hat{y}^m \left((1 - \beta) \frac{(1 - \gamma^*)^\alpha_{HH} + \alpha_{HL}}{\alpha_{LL}}\right) & Y^w_{LL} &= \hat{y}^w \left((1 - \beta) \frac{\gamma^* \alpha_{HH} + \alpha_{HL}}{\alpha_{LL}}\right) \\
Y^m_{LH} &= \hat{y}^m \left((1 - \beta) \frac{\gamma^* \alpha_{HH}}{\alpha_{HL}}\right) & Y^w_{LH} &= \hat{y}^w \left((1 - \beta) \frac{(1 - \gamma^*) \alpha_{HH}}{\alpha_{HL}}\right)
\end{align*}

for $\gamma^* \in (0, 1)$ satisfying

\begin{align*}
\phi^m \left((1 - \beta) \frac{(1 - \gamma^*)^\alpha_{HH} + \alpha_{HL}}{\alpha_{LL}}\right) &= \phi^m \left((1 - \beta) \frac{\gamma^* \alpha_{HH}}{\alpha_{HL}}\right) \\
&= \phi^w \left((1 - \beta) \frac{(1 - \gamma^*) \alpha_{HH}}{\alpha_{HL}}\right).
\end{align*}

It holds that $Y^m_{LL} \leq Y^m_{LH} = Y^m_{HH}$, and $Y^w_{LL} \leq Y^w_{LH} \leq Y^w_{HH} = Y^w_{HH}$. The IC constraints \{IC$_{HH,HL}$, IC$_{HH,LH}$, IC$_{LL,HH}$, IC$_{HL,HH}$\} are binding.

---

8 Abusing some notation to clarify the relationship between the papers, the function $v(\cdot)$ in Armstrong and Rochet (1999) is set to 0 in my model. The function $u^{i\theta}(\cdot)$ in Armstrong and Rochet (1999) corresponds to $y - C^i(y | \theta^i)$, which I later denote $v^i(y | \theta^i)$. 
Strong Positive Correlation. If $\rho > \frac{\alpha_{HH} \alpha_{HL}}{\alpha_{LL}}$, then the optimal policy induces output levels

$$Y_{HH}^m = Y_{HL}^m = Y_{H}^m = Y_{HH}^w = Y_{LH}^w = Y_{H}^w$$

$$Y_{LL}^m = Y_{LH}^m = \tilde{y}^m \left(1 - \beta \right) \frac{\alpha_{HH} + \alpha_{HL}}{\alpha_{LL} + \alpha_{HL}}$$

$$Y_{LL}^w = Y_{LH}^w = \tilde{y}^w \left(1 - \beta \right) \frac{\alpha_{HH} + \alpha_{HL}}{\alpha_{LL} + \alpha_{HL}}.$$ 

It holds that $Y_{LL}^m = Y_{LH}^m \leq Y_{HL}^m = Y_{HH}^m$, and $Y_{LL}^w = Y_{LH}^w \leq Y_{HL}^w = Y_{HH}^w$. The IC constraints $\{IC_{HH,HL}, IC_{HH,LH}, IC_{HH,LL}, IC_{LH,LL}, IC_{HL,LL}\}$ are binding.

From the theorem, we see that high individuals always produce an efficient level of output which is independent of their spouse’s type. This result replicates the property of “no distortion at the top” familiar from many studies of optimal taxation. Only downward incentive constraints bind, so there is no reason for the planner to distort the high type choices.

Low individuals, on the other hand, earn less than the efficient level. Under Weak Positive Correlation, a low individual married to a low spouse earns even less than one who is married to a high type spouse. Under Strong Positive Correlation, a low individual’s earnings do not depend on his or her spouse’s type.

As Armstrong and Rochet discuss, the key to the proof is to determine which of the IC constraints bind. In the case of Strong Positive Correlation, all of the “downward” constraints $\{IC_{HH,HL}, IC_{HH,LH}, IC_{HH,LL}, IC_{LH,LL}, IC_{HL,LL}\}$ are binding. With Weak Positive Correlation, four out of these five constraints bind; the “diagonal” constraint $IC_{HH,LL}$ is lax. In either case the government does not need to consider “upward” constraints preventing low types from pretending to be high types, or “transverse” constraints preventing one mixed type from pretending to be the opposite mixed type.\(^9\)

### III. Marginal Tax Rates and Jointness

#### A. Marginal Tax Rates

There is a one-to-one relationship between output levels and marginal tax rates. Higher outputs correspond to lower tax rates. So the results from Theorem 1 regarding output levels under the optimal tax code directly imply the following properties of marginal tax rates:

**PROPOSITION 1:** For $\rho \geq 0$, the optimal marginal tax rates satisfy the following properties:

(i) **High types:** zero marginal tax rate.
   
   If $\theta^i = H$, then $t^i_0 = 0$. 

\(^9\) While I only discuss the case of nonnegative correlation, Armstrong and Rochet (1999) cover general distributions. For costs that are close to symmetric, for instance, the solution would be the same as under Weak Positive Correlation even with disassortative mating.
**Proposition 2:** Under Weak Positive Correlation, the optimal tax code has negative jointness. Under Strong Positive Correlation, the optimal tax code has zero jointness.

Recall that the models of Kleven, Kreiner, and Saez (2007, 2009) derive negative jointness results under the assumption of independent types. In the current model we see that negative jointness extends to any Weak Positive Correlation, but not to Strong Positive Correlation. When both genders are split evenly between high and low types, for example, we have negative jointness when the correlation coefficient of men and women’s types is below $1/3$, and zero jointness for higher correlations.

To understand why the jointness of the tax code depends on assortative mating, it is helpful to consider a variational argument of the sort in Kleven, Kreiner, and Saez (2007, 2009). For simplicity, consider an economy with Symmetric Genders. (The proposition does not depend on this assumption.) Start with the best possible zero-jointness tax schedule. We can perturb the tax code towards negative jointness by having the low types in low couples work slightly less, and the low types in mixed couples work slightly more. In particular, the low types in low couples both reduce their output by $1/\alpha_{LL}$ infinitesimal units while low types in $HL$ and $LH$ couples each increase output by $1/\alpha_{HL} = 1/\alpha_{LH}$ units. This perturbation yields the same aggregate costs of effort across all workers in the economy, and the same aggregate output. So
if we combine this output perturbation with a redistribution of money across types, social welfare will have increased if and only if low-low utility has gone up. Is it possible to redistribute money in a budget-balanced way that makes low-low types better off, while also maintaining incentive compatibility of the perturbed tax code? I seek to explain why this is possible only for relatively low levels of correlation.

The limit on redistribution comes from the incentive constraints. At the best zero-jointness schedule, the downward incentive constraints all bind. After perturbing outputs towards negative jointness and transferring enough to low-low couples to make them better off, we need to make sure that no couples now want to mimic a lower type. The higher is the LL after-tax income, the higher must be HL and LH couples’ after-tax income in order to keep them from pretending to be LL; and the higher must be HH couples incomes as well, to keep them from pretending to be HL or LH. Say that the minimum possible change in after-tax income to type $\theta$ which both makes the low-low types weakly better off, and which continues to satisfy incentive compatibility for other types, is given by $\delta_\theta$. Such transfers can be made in a budget-balanced manner if $\sum_\theta \alpha_\theta \delta_\theta \leq 0$; when the inequality is strict, there is money left over to make the low-low types (and thus social welfare) strictly better off.

In order to calculate the transfers $\delta_\theta$, we need to know the effort costs imposed by the output perturbation. At the original low type output level, say that low types pay a marginal utility cost of $r > 0$ for each additional dollar of output. If high types were to deviate to this level of output, they would face a marginal cost of $s$ per dollar of extra output, with $0 < s < r$ because high types have a lower cost of effort.

After the perturbation, individuals in low-low couples work less than before; the effort cost to each has gone down by $r/\alpha_{LL}$. So we can make the couples better off even if we reduce their after-tax income. We must only transfer at least $\delta_{LL} = -2r/\alpha_{LL}$ dollars to each low-low couple. Look now at mixed couples, which prior to the perturbation were indifferent to deviating to LL; without loss, focus on the LH couples. After the output perturbation and the transfers, the payoff of continuing to take the mixed outputs has changed by $\delta_{LH} - r/\alpha_{LH}$; adding up the change in after-tax income and the extra costs of effort. On the other hand, the payoff of pretending to be LL goes up by $\delta_{LL} + r/\alpha_{LL} + s/\alpha_{LL}$; the low type in the mixed couple saves $r$ per unit reduction in output, and the high type saves $s$. To maintain incentive compatibility, the former must outweigh the latter. That implies a minimum transfer to mixed types of $\delta_{LH} = \delta_{LL} + r/\alpha_{LH} + s/\alpha_{LL} + r/\alpha_{LH} = r/\alpha_{LH} - (r - s)/\alpha_{LL}$. Similarly, in order to keep the high-high couples from deviating to the new mixed outputs, they must be given an extra payment of $\delta_{HH} = \delta_{LH} - s/\alpha_{LH} = (r - s)/(1/\alpha_{LH} - 1/\alpha_{LL})$.

The net transfer to all of the HL, LH, and LL couples is $\alpha_{LL} \delta_{LL} + 2\alpha_{LH} \delta_{LH} = -2(r - s)\alpha_{LH}/\alpha_{LL}$. This is negative, which means we get a surplus from these three types of couples. However, as we add assortative mating to the economy, this surplus goes down; the $\alpha_{LH}$ in the numerator decreases while the denominator $\alpha_{LL}$ increases. On the other hand, the net transfer to HH couples is $\alpha_{HH} \delta_{HH} = (r - s)(\alpha_{HH}/\alpha_{LH} - \alpha_{HH}/\alpha_{LL})$. As we add correlation to the economy and reduce the number of mixed couples, the positive term $\alpha_{HH}/\alpha_{LH}$ blows up. We can try to use the surplus from the couples with low types to subsidize this deficit from high-high couples, but the cross subsidization becomes more difficult as we add
assortative mating. The surplus declines while the deficit increases. So this perturbation can only be made in a budget balanced manner for low enough levels of assortative mating. Indeed, this perturbation is budget balanced if and only if we are in the region of Weak Positive Correlation.

A similar variational argument also clarifies why the tax code becomes zero joint rather than positively joint as we move into the range of Strong Positive Correlation. One might have supposed that under large correlations, we could improve welfare with a perturbation towards positive jointness, exactly the reverse of the one described above. The reason we cannot do so is due to the diagonal HH-LL incentive constraint, which binds at the optimal zero-joint tax schedule under Strong Positive Correlation. I ignored the HH-LL constraint in the argument above because the perturbation towards negative jointness relaxes that constraint: having LL types work less hard and keep less money makes a deviation of HH couples to LL outputs less appealing. If we were to run the perturbation in reverse, however, the HH-LL constraint would tighten. Maintaining incentive compatibility for HH types would require too large a subsidy to be able to satisfy budget balance while improving LL utilities.10

C. Additional Comparative Statics

The tax code switches from negative jointness to zero jointness at the cutoff of Weak and Strong Positive Correlation, but there is no discontinuity in the tax code at this point. Tax rates all move continuously with the degree of assortative mating. See Figure 1 for an illustration of how tax rates and corresponding outputs vary with the degree of assortative mating. For the numerical analysis in the figure I take genders to be symmetric; I let $c_i(y|H) = y^2/2$ and $c_i(y|L) = y^2$; I let the redistribution parameter be $\lambda = 1/2$; and I suppose that both genders are split half-half between high and low types ($q = 1/2$). The level of assortative mating is parametrized by $\rho$, where mating is uncorrelated at $\rho = 0$ and is fully assortative at $\rho = 1/4$.

The figure confirms that in the case of Weak Positive Correlation, the tax rate on low types in matched couples is higher than that for low types in mixed couples. For Strong Positive Correlation, though, the tax rates do not depend on partner type. This gives us negative jointness for low correlations, and zero jointness for high correlations (Proposition 2).

Moreover, Figure 1 suggests a sense in which the degree of negative jointness is attenuated as the level of assortative mating increases. Over the range of Weak Positive Correlation, the tax rate on low individuals with low partners falls in $\rho$ as the tax rate on low individuals with high partners rises. So we have the greatest degree of negative jointness when the correlation parameter is 0, and we move continuously to a tax code with zero jointness as the correlation increases. Proposition 3

---

10 On a technical level, it follows from simple algebra that if all of the downward incentive constraints bind, including the HH-LL one, then low output must be independent of partner type. The part that is difficult to show, which follows from the analysis of Armstrong and Rochet (1999)—see Theorem 1, above—is that all of these constraints bind at the optimal tax schedule under Strong Positive Correlation.
formalizes this attenuation of negative jointness with respect to the level of assortative mating for the case of Symmetric Genders.

PROPOSITION 3: Impose Assumption 1 (Symmetric Genders), and fix the proportion of high individuals $q \in (0, 1)$ in each gender while letting the joint type distribution vary with the assortative mating parameter $\rho$. Under Weak Positive Correlation:

(i) The tax rates on low individuals in low-low couples, $t_{LL}^m$ and $t_{LL}^w$, weakly decrease in $\rho$.

(ii) The tax rates on low individuals in mixed couples, $t_{HL}^m$ and $t_{HL}^w$, weakly increase in $\rho$.

Changes in output levels $Y^i_\theta$ have the opposite sign of changes in tax rates $t^i_\theta$.

PROOF:
See the Mathematical Appendix.

In the numerical example of Figure 1 we also see that, in the range of Strong Positive Correlation, the optimal tax rate on low types rises in the degree of assortative mating. Proposition 4, below, confirms that this holds in general. This result, unlike that of Proposition 3, does not depend on an assumption of Symmetric Genders.

PROPOSITION 4: Fix a distribution of high and low individuals for each gender, $q^m$ and $q^w$, while letting the joint type distribution vary with the assortative mating parameter $\rho$. Under Strong Positive Correlation:

(i) The tax rate on low men in any type of couple, $t_{LL}^m = t_{LH}^m$, weakly increases in $\rho$. 

![Figure 1. Marginal Tax Rates and Outputs for Low Individuals in Low-Low and Mixed Couples](image-url)
(ii) The tax rate on low women in any type of couple, \( t_{LH}^w = t_{HL}^w \), weakly increases in \( \rho \).

Changes in output levels \( Y_\theta^i \) have the opposite sign of changes in tax rates \( t_\theta^i \).

PROOF: See the Mathematical Appendix.

IV. Tax Levels

In this section, I turn attention to the tax levels (net transfers) \( t_\theta \) paid by each type of couple.

To simplify notation going forward, it is convenient to define a new expression \( v_i(y|\theta^i) \equiv y - c_i(y|\theta^i) \) denoting the pretax surplus generated by an individual of gender \( i \) and type \( \theta^i \) who produces \( y \) units of output. It holds that \( v_i(y|\theta^i) \) increases in \( y \) for \( y \leq y_{\theta}^i \m^* \). In this new notation, the IC constraint \( IC_{0,\theta'} \) can be written as

\[
\sum_i v_i(Y_\theta^i|\theta^i) - t_\theta \geq \sum_i v_i(Y_{\theta'}^i|\theta^i) - t_{\theta'}, \quad \text{or equivalently}
\]

\[
(IC_{0,\theta'}) \quad t_\theta - t_{\theta'} \leq \sum_{i=m,w} v_i(Y_\theta^i|\theta^i) - \sum_{i=m,w} v_i(Y_{\theta'}^i|\theta^i).
\]

Under either Weak or Strong Positive correlation, the IC constraints \( IC_{HH,HL}, IC_{HH,LL}, IC_{HL,HL}, \) and \( IC_{HL,LL} \) all hold with equality. And in the case of Strong Positive Correlation, we have a zero joint tax code; the output \( y_\theta^i \) is independent of \( \theta^{-i} \). So for Strong Positive Correlation these four IC constraints simplify to \( \tau_{HH} - \tau_{HL} = \tau_{LH} - \tau_{LL} = v^w(Y_{LH}^w|H) - v^w(Y_{LL}^w|H) \), and \( \tau_{HH} - \tau_{LH} = \tau_{HL} - \tau_{LL} = v^m(Y_{HL}^m|H) - v^m(Y_{LL}^m|H) \). The former chain of equalities tells us that the extra tax on a woman from working at the high output level rather than the low output level does not depend on her husband’s type. The latter chain tells us that the same holds for a man with respect to his wife’s type. These two chains of inequalities combined with the previous observation of zero jointness imply that the tax code is separable. That is, it can be implemented by an individualized tax code in which one’s taxes are independent of his or her partner’s income.

PROPOSITION 5: Under Strong Positive Correlation, the optimal tax code is separable.

We can also use these incentive compatibility constraints to compare tax levels across high-high, mixed, and low-low types. As one might expect, the high-high types pay a net tax and the low-low types receive a subsidy. The tax levels of mixed couples are of ambiguous sign.

PROPOSITION 6: High-high couples pay more taxes than mixed couples, and mixed couples pay more taxes than low-low couples: \( \tau_{HH} \geq \max\{\tau_{HL}, \tau_{LH}\} \) and \( \tau_{LL} \leq \min\{\tau_{LH}, \tau_{HL}\} \). Moreover, high-high couples pay a net tax while low-low couples receive a net subsidy: \( \tau_{HH} \geq 0, \) and \( \tau_{LH} \leq 0 \).
PROOF:
See the Mathematical Appendix.

The exact tax levels can be solved for by taking any three of the four binding non-diagonal downward IC constraints (one is redundant) and combining them with the budget balance condition \( \sum_\theta \alpha_\theta \tau_\theta = 0 \) to get four independent linear equations for the four terms \( \tau_\theta \). The interested reader can find the Symmetric Genders tax levels worked out in Lemma 1 of the Mathematical Appendix. Figure 2 illustrates net tax levels paid by each type of couple for Symmetric Genders with \( C^i(y|H) = y^2/2 \), \( C^i(y|L) = y^2 \), and \( \lambda = 1/2 \). Panel A shows the case of few high types, with \( q = 1/4 \); panel B shows many high types, \( q = 3/4 \). The figure confirms that high-high couples pay positive taxes, and low-low couples receive net subsidies. (To give context for the magnitudes of the tax levels in the graph, low individuals each output a little under 0.5 units, and high individuals output 1 unit.) In the examples, mixed couples pay positive taxes when there are few high types, but receive subsidies when there are many high types. Some comparative statics on the tax levels under Strong Positive Correlation are discussed in the following section.

V. Further Characterizations under Strong Positive Correlation

The separability of the optimal tax code under Strong Positive Correlation makes it relatively simple to take additional comparative statics. In this section, I show how the tax code and the utility levels of each type of couple vary with the degree of assortative mating.

Recall that the utility of a couple is the after-tax income minus the individual effort costs. When the tax code is separable, as it is under Strong Positive Correlation, we can allocate this net tax bill across the two individuals as \( \tau_\theta = \tau_{\theta m} + \tau_{\theta w} \). Thus, we can think of a couple’s utility as the sum of the two spouses’ implied individual utilities. Write these implied utilities as \( U_{\theta i} = Y_{\theta i} - C^i(Y_{\theta i} | \theta^i) - \tau_{\theta i} \), so that \( U_\theta = U_{\theta m} + U_{\theta w} \).
Under Strong Positive Correlation the optimal tax code can be found in the restricted class of separable tax codes. Optimizing over tax codes in this class is equivalent to performing two independent maximizations, over the men’s tax code and the women’s tax code.\footnote{Without loss of generality, there is no subsidization across genders; we can take budget balance to be satisfied separately for men and for women.} Writing $q^i$ as the proportion of high types in gender $i$, the government solves the following program for gender $i$:\footnote{The only binding incentive compatibility constraint is the $H$ to $L$ one. I also omit the denominator $\alpha_{LL} + \lambda (1 - q^i)$ on the objective function, because it is decision irrelevant.}

\[
\begin{align*}
\max & \left\{ y^i, \tau^i, \ U^i \right\} \left( \alpha_{LL} + \lambda \left( (1 - q^i) - \alpha_{LL} \right) \right) U^i_L + \lambda q^i U^i_H \\
\text{s.t.} & \quad U^i_L = Y^i_L - C^i(Y^i_L \mid \theta^i) - \tau^i \\
\left( \text{IC}_{H,L} \right) & \quad U^i_H \geq U^i_L + C^H(Y^i_L \mid L) - C^i(Y^i_L \mid H) \\
\left( \text{BB} \right) & \quad (1 - q^i) \tau^i_L + q^i \tau^i_H = 0.
\end{align*}
\]

In the objective function here, the weight placed on high type utility is $\lambda$ times the number of high types. The weight on low type utility is $\lambda$ times the number of low types in mixed couples, plus 1 times the number of low types in low-low couples.

For a fixed proportion $q^i$ of high types in the population, the level of assortative mating $\rho$ affects the maximization program only through the $\alpha_{LL}$ term in the objective function. Every other term in the objective function and in the constraints is independent of $\rho$. So as we increase the level of assortative mating, we put a higher effective Pareto weight $\lambda q^i$ on $Y^i_H$ while every other term remains fixed.\footnote{The analysis is not so clear in the range of Weak Positive Correlation. In numerical examples, for instance, low-low utility can nonmonotonically decrease and then increase in $\rho$ while high-high utility monotonically decreases.} It is immediate from this observation that $U^i_L$ must weakly increase with $\rho$ while $U^i_H$ weakly decreases.

Returning to the analysis of couples, it follows that $U_{LL} = U^m_L + U^w_L$ is increasing in $\rho$ while $U_{HH} = U^m_H + U^w_H$ is decreasing. That is, as the assortative mating increases in the range of Strong Positive Correlation, the utility of each low-low couple improves while the utility of high-high couple goes down. Mixed couples average out these two effects, and so their utility change is ambiguous; numerical examples show that it can go in either direction.

In other words, an increase in assortative mating will have two opposing effects on utility inequality across couples. The direct effect is to amplify inequality by increasing the number of low-low and high-high couples relative to mixed ones, as discussed in the Introduction. There is another effect, however, due to the reoptimization of the tax code. This second effect pushes towards equality: the utilities of the highest and lowest couples move closer together.\footnote{The analysis is not so clear in the range of Weak Positive Correlation. In numerical examples, for instance, low-low utility can nonmonotonically decrease and then increase in $\rho$ while high-high utility monotonically decreases.}

The utility changes also imply comparative statics in tax levels with $\rho$. We know that high type effort will always be efficient at the optimum: $Y^i_H = y^i_H$. So if high individuals get lower utility in a reoptimized tax code, it must be because they pay...
higher taxes. And by budget balance, if high individuals pay a higher tax then low individuals must face lower taxes—i.e., receive a higher subsidy. So low-low couples pay decreasing net tax levels $\tau_{LL}$, and high-high couples pay increasing tax levels $\tau_{HH}$. Mixed couples once again face an average of these two changes, and their tax bill can go in either direction.$^{14}$

Collecting these arguments:

PROPOSITION 7: Fix a distribution of high and low individuals for each gender, $q^m$ and $q^w$, while letting the joint type distribution vary with the assortative mating parameter $\rho$. Under Strong Positive Correlation:

(i) High-high utility $U_{HH}$ is weakly decreasing in $\rho$, while low-low utility $U_{LL}$ is weakly increasing.

(ii) High-high tax levels $\tau_{HH}$ are weakly increasing in $\rho$, while low-low tax levels $\tau_{LL}$ are weakly decreasing.

VI. Conclusion

This paper presents a model of the optimal taxation of couples in the presence of assortative mating. I characterize the optimal tax schedule for general effort cost functions and joint distributions of types, allowing for potential asymmetries across genders. In order to solve the model, however, I make some strong assumptions that I would like to highlight here.

Most prominently, I assume that there are only two possible types for each individual. This assumption was made for tractability, and to allow for the simplest and clearest possible exploration of assortative mating. However, binary models of taxation can be misleading. The standard result of zero marginal tax rates for the single highest type, for instance, takes on a very different character when there are only two types than when types are continuous. Indeed, in the current model, this fact of no marginal taxation at the top meant that the jointness results were driven by changes in tax rates on low types. That said, the two-by-two model is rich enough to replicate the Kleven, Kreiner, and Saez (2007, 2009) results of negative jointness in the absence of assortative mating. The model also suggests an intuition for why such jointness may attenuate with correlated types. When we add negative jointness to the tax code, we have to reduce tax levels on couples with two high earners in order to keep them working hard. These tax reductions become increasingly unaffordable as the correlation of spousal earnings goes up.

The model also assumed a particular form for the government’s redistributive preference: the government overweights the utility of the low-low couples, but cares equally about mixed and high-high couples. I required this social welfare function in

$^{14}$Under Symmetric Genders as well as Strong Positive Correlation, Lemma 1 in the Mathematical Appendix implies that mixed tax levels are $\tau_{HL} = \tau_{LH} = (1 - 2q)(v^w(Y_{HH}^w|H) - v^w(Y_{LL}^w|H))$. From this equation, we see that mixed couples face a nonnegative tax level which increases in $\rho$ when high individuals are rare in the population ($q < \frac{1}{2}$). When high individuals are common ($q > \frac{1}{2}$), the tax level is weakly negative and is decreasing in $\rho$. 
order to solve for the optimal tax schedule, as a step towards transforming my mathematical problem into that of Armstrong and Rochet (1999). It is natural to wonder how the tax code would vary under alternative welfare functions. For instance, the government might care most about low-low couples, but also care more about mixed than high-high couples.

The variational argument following Proposition 2 suggests that this modification of the social welfare function would make negative jointness somewhat less attractive. I consider a perturbation of the tax code from separability to negative jointness, and observe that the perturbation will be welfare improving as long as low-low couples are made better off. This is feasible for sufficiently low levels of assortative mating, i.e., Weak Positive Correlation; at the point where we switch to Strong Positive Correlation, low-low couples can only be paid enough to make them exactly as well off as before. At this cutoff correlation, the perturbation harms mixed couples and benefits high-high couples by exactly offsetting amounts. Under the proposed alternative social welfare function, though, the harms and benefits would not cancel out. Transferring an equal amount of utility from a mixed couple to a high-high couple would be a net welfare negative. Higher payments to low-low or mixed couples would be required in order to make the perturbation welfare improving, and would only be affordable at lower levels of correlation. Hence, under the alternative social welfare function we might expect the same basic features of the tax code: negatively joint at low levels of assortative mating, separable at higher levels. But the negatively jointness would likely hold over a smaller range of correlations.

**Mathematical Appendix**

**PROOF OF PROPOSITION 3:**

From Theorem 1, to show that tax rate \( t_\theta \) is weakly increasing (decreasing) with \( \rho \), it suffices to show that \( \zeta \) is weakly increasing (decreasing), where \( \zeta \) is the respective argument of \( Y_\theta = \hat{y}^i(\zeta) \):

\[
\begin{align*}
\zeta_{LL}^m &= (1 - \beta) \left(1 - \gamma^*\right) \alpha_{HH} + \alpha_{HL} \, \zeta_{LL}^w &= (1 - \beta) \left(1 - \gamma^*\right) \alpha_{HH} + \alpha_{HL} \\
\zeta_{LH}^m &= (1 - \beta) \gamma^* \alpha_{HH} \, \zeta_{HL}^w &= (1 - \beta) \left(1 - \gamma^*\right) \alpha_{HH} \, \zeta_{HL}^w &= (1 - \beta) \left(1 - \gamma^*\right) \alpha_{HH} \, \zeta_{HL}^w &= (1 - \beta) \left(1 - \gamma^*\right) \alpha_{HH} \\
\end{align*}
\]

where \( \gamma^* \in (0, 1) \) is defined as the solution to

\[
\gamma^* = \frac{1}{\alpha_{HH}} \left( \alpha_{HL} - \alpha_{HH} \right)
\]

15The logic of Kleven, Kreiner, and Saez (2007, 2009) provides a separate intuition pointing in the same direction. As I discuss in footnote 6, their negative jointness result requires convexity of the derivative of their redistribution function \( \Psi \); if \( \Psi' \) were concave, the tax code would be positively joint. Reducing the welfare weight on high-high types effectively makes my \( \Psi' \) less convex, which makes negative jointness less appealing.
\[
\phi_m'(1 - \beta) \left( \frac{(1 - \gamma^*) \alpha_{HH} + \alpha_{HL}}{\alpha_{LL}} \right) - \phi_m'(1 - \beta) \left( \frac{\gamma^* \alpha_{HH}}{\alpha_{HL}} \right)
= \phi_w'(1 - \beta) \left( \frac{\gamma^* \alpha_{HH} + \alpha_{HL}}{\alpha_{LL}} \right) - \phi_w'(1 - \beta) \left( \frac{(1 - \gamma^*) \alpha_{HH}}{\alpha_{HL}} \right).
\]

The function \( \phi_m'(\cdot) \) is equal to \( \phi_w'(\cdot) \) when cost functions are identical across genders. Under the assumption of Symmetric Genders, we see that this equation is solved with \( \gamma^* = 1 - \gamma^* \), i.e., \( \gamma^* = 1/2 \). Plugging in \( \gamma^* = 1/2 \) and \( \beta = \lambda/(\alpha_{LL} + \lambda(1 - \alpha_{LL})) \) into the \( \zeta \) expressions and then rewriting all expressions in terms of \( q \) and \( \rho \), we get

\[
\zeta_{LL}^m = \zeta_{LL}^w = \frac{(1 - 2q + q^2 + \rho)(1 - \lambda)}{1 - 2q + q^2 + \rho + \lambda(2q - q^2 - \rho)} \frac{1}{2} \left( q^2 + \rho + (q - q^2 - \rho) \right)
\]

\[
\zeta_{LH}^m = \zeta_{HL}^w = \frac{(1 - 2q + q^2 + \rho)(1 - \lambda)}{1 - 2q + q^2 + \rho + \lambda(2q - q^2 - \rho)} \frac{1}{2} \left( q^2 + \rho \right)
\]

Evaluating these derivatives with respect to \( \rho \) gives the desired comparative statics results.

**PROOF OF PROPOSITION 4:**

I will prove the results for tax rates on men; results for women follow identically.

From Proposition 1, we want to show that \( (1 - \beta)(\alpha_{HH} + \alpha_{HL})/(\alpha_{LL} + \alpha_{HL}) \) is increasing in the level of assortative mating. The term \( (\alpha_{HH} + \alpha_{HL})/(\alpha_{LL} + \alpha_{HL}) \) is independent of the level of assortative mating \( \rho \)—the numerator is the number of high types of men, and the denominator is the number of low types for men. So it is sufficient to show that \( (1 - \beta) \) is increasing in correlation, i.e., \( \beta \) is decreasing in correlation.

Recall that \( \beta \) was defined as \( \lambda/(\alpha_{LL} + \lambda(1 - \alpha_{LL})) \). So \( \beta \) declines in \( \alpha_{LL} \), as long as \( 0 \leq \lambda \leq 1 \). And \( \alpha_{LL} \) increases as we increase the correlation of types \( \rho \).

**PROOF OF PROPOSITION 6:**

Let us first expand out the four always-binding IC constraints in terms of \( v^i \) notation. The fact that there is no distortion at the top \( Y^i_0 = Y^i_{\theta^i} \) if \( \theta^i = H \) allows us to cancel out some terms:

\[
(IC_{HH,HL}) \quad \tau_{HH} - \tau_{HL} = v^w(Y^w_{HH} | H) - v^w(Y^w_{HL} | H)
\]

\[
(IC_{HL,LL}) \quad \tau_{HL} - \tau_{LL} = (v^m(Y^m_{HL} | H) - v^m(Y^m_{LL} | H)) + (v^w(Y^w_{HL} | L) - v^w(Y^w_{LL} | L))
\]
\[(IC_{HH,LL}) \quad \tau_{HH} - \tau_{LL} = v^m(Y_{HH}^m | H) - v^m(Y_{LL}^m | H)\]

\[(IC_{LH,LL}) \quad \tau_{LH} - \tau_{LL} = (v^w(Y_{LH}^w | H) - v^w(Y_{LL}^w | H)) + (v^m(Y_{LH}^m | L) - v^m(Y_{LL}^m | L)).\]

We can see from \(IC_{HH,HL}\) that \(\tau_{HH} \geq \tau_{HL}\), because \(Y_{HH}^w \geq Y_{HL}^w\) and \(v^w(y | H)\) is increasing over the relevant range—indeed, \(v^w(y | H)\) is maximized at \(y = Y_{HH}^w\). Likewise, \(\tau_{HH} \geq \tau_{LH}\) from \(IC_{HH,LH}\). Similar arguments also show that \(\tau_{HL} \geq \tau_{LL}\) and \(\tau_{LH} \geq \tau_{LL}\); each of the parenthetical terms on the right-hand side of the above expressions is weakly positive, with \((v^w(Y_{HL}^w | L) - v^w(Y_{LL}^w | L))\) and \((v^w(Y_{LH}^w | H) - v^w(Y_{LL}^w | H))\) equal to 0 in the case of Strong Positive Correlation.

By budget balance, the highest tax \(\tau_{HH}\) must be greater than or equal to 0, and the lowest tax \(\tau_{LL}\) must be less than or equal to 0.

**LEMMA 1:** Under Assumption 1 (Symmetric Genders), tax rates are as follows (with \(Y_{0}^i\) values determined by Theorem 1).\(^{16}\)

\[
\tau_{HH} = (1 - 2q + q^2 + \rho)(v^w(Y_{HH}^w | H) - v^w(Y_{LL}^w | H)) + (v^w(Y_{HL}^w | L) - v^w(Y_{LL}^w | L)) \\
+ (1 - q^2 - \rho)(v^w(Y_{HH}^w | H) - v^w(Y_{HL}^w | H)) \\
= \alpha_{LL}\left((v^w(Y_{HH}^w | H) - v^w(Y_{LL}^w | H)) + (v^w(Y_{HL}^w | L) - v^w(Y_{LL}^w | L))\right) \\
+ (1 - \alpha_{HH})(v^w(Y_{HH}^w | H) - v^w(Y_{HL}^w | H))
\]

\[
\tau_{HL} = \tau_{LH} = (1 - 2q + q^2 + \rho)(v^w(Y_{HH}^w | H) - v^w(Y_{LL}^w | H)) \\
+ (v^w(Y_{HL}^w | L) - v^w(Y_{LL}^w | L)) \\
- (q^2 + \rho)(v^w(Y_{HH}^w | H) - v^w(Y_{HL}^w | H)) \\
= \alpha_{LL}\left((v^w(Y_{HH}^w | H) - v^w(Y_{LL}^w | H)) \\
+ (v^w(Y_{HL}^w | L) - v^w(Y_{LL}^w | L))\right) \\
- \alpha_{HH}(v^w(Y_{HH}^w | H) - v^w(Y_{HL}^w | H))\]

\(^{16}\)I write the expressions in terms of \(v^w\) and \(Y_{0}^i\), but of course they could be equally written in terms of \(m\) rather than \(w\). Furthermore, I replace all occurrences of \(Y_{LL}^w\) with the equal value of \(Y_{HH}^w\). Recall from Assumption 1 that \(q\) is the proportion of high individuals and \(\rho\) is the covariance term. In the expressions, note that all of the parenthetically sized \((v^w(y | \theta^i) - v^w(y | \theta^i))\) terms are weakly positive.
\[ \tau_{LL} = (2q - q^2 - \rho)((v^w(Y^w_{HH} | H) - v^w(Y^w_{LL} | H)) \\
+ (v^w(Y^w_{HL} | L) - v^w(Y^w_{LL} | L))) \\
- (q^2 + \rho)(v^w(Y^w_{HH} | H) - v^w(Y^w_{HL} | H)) \\
= -(1 - \alpha_{LL}) \left( (v^w(Y^w_{HH} | H) - v^w(Y^w_{LL} | H)) + (v^w(Y^w_{HL} | L) - v^w(Y^w_{LL} | L)) \right) \\
- \alpha_{HH}(v^w(Y^w_{HH} | H) - v^w(Y^w_{HL} | H)). \]

**PROOF OF LEMMA 1:**
These formulas result from solving the three linear equations of \((IC_{HH,HL})\), \((IC_{HL,LL})\), and \((BB)\) for the three unknown tax rates. (The IC constraints can be found expanded out in terms of \(v^i\) notation in the proof of Proposition 6.)

**REFERENCES**


