Online Appendix

A. Non-sex consumption

In this section, we consider model extensions with a non-sex consumption good. This introduces, in principle, the possibility that a "tax" on prostitution shifts male expenditure away from both markets of sexual interaction, which could lead to a decrease in the price of (non-reproductive) sex, p_s, such that no over-compensation effect arises.

This is not the case under standard assumptions on marginal utilities. Under optimal consumption schedules, taxing prostitution (weakly) increases marriage expenditure p_m and hence does not reduce the price of sex p_s in case A of our model, in which p_s moves in lockstep with p_m through the women's indifference condition.

Assumptions. Consider a decriminalized market without trafficking. The populations of men and women each have unit mass.

A man allocates his budget y across three consumption goods: He spends p_m on (the quality of his) marriage, if married; $x_s p_s$ on sex from prostitutes; and x_o non-sex consumption. His budget constraint is $p_m + x_s p_s + x_o = y$ where p_s is the price of sex in the prostitution market. It is important to note that x_o is consumption "outside" of marriage, that is, expenditure that does not increase marital surplus or consumption.

A man's utility is given by $U = S(x_s) - p_s x_s + M(p_m) + O(x_o)$. We make the standard assumption that the marginal utilities are positive but decreasing: $S', M', O' \geq 0$ and $S'', M'', O'' \leq 0$. We intentionally keep (marginal) utility of each good independent; there is no complementarity or substitutability directly built into consumption preferences that could drive a co-movement in prices.

The assumptions for women are as in the original model: a woman's wage in the regular labor market is w; occupational hazards impose disutility $-h$ on prostitutes; the number of voluntary prostitutes (and thus unmarried men) is denoted by n.

In what follows, we restrict our attention to the case of interior equilibria.

Interior equilibrium. In an interior equilibrium, these conditions hold: A married man's consumption equalizes marginal utility across all three goods,

$$S'(x_s^m) - p_s = M'(p_m) = O'(x_o^m) \tag{1}$$

subject to his budget constraint $p_m + x_s^m p_s + x_o^m = y$.

An unmarried man's consumption equalizes marginal utility across non-sex consumption and the consumption of non-reproductive sex,

$$S'(x_s^u) - p_s = O'(x_o^u) \tag{2}$$

subject to his budget constraint $x_s^u p_s + x_o^u = y$. Superscripts m and u indicate "married" and "unmarried."

A woman's indifference condition is, as in the original model,

$$sp_s = w + \Delta. \tag{3}$$

where $\Delta = h + p_m$ is a compensating differential.

Last, market clearing implies

$$nsp_s = y - (1 - n)(p_m + x_o^m) - nx_o^u, \tag{4}$$

that is, the total income of prostitutes equals the men's total budget less married men's spending on marriage and non-sex consumption and less unmarried men's spending on non-sex consumption.

A "tax" on demand. Considering interior equilibrium, we are interested in the effect of a change in function $S(.)$ that lowers men's (marginal) utility from buying non-reproductive sex on the price of sex p_s. There are two countervailing effects:

- The *classic* effect on men's demand, per (1), is to shift male spending toward the other goods, i.e., to increase p_m or x_o (or both) and decrease consumption of non-reproductive sex x_s and the price of sex p_s until marginal utilities are again equalized. Without the feedback effect (discussed next), this usually implies a *higher* p_m and a *lower* p_s.
- The women's labor choice creates the *feedback* effect that changes in p_m alter the compensating differential Δ in (3) and hence the price of sex p_s. Indeed, (3) implies that p_s and p_m ultimately move in the *same* direction in an interior equilibrium.

In principle, p_s can hence decrease or increase, or stay unchanged. Both of the first two outcomes have counterintuitive aspects due to the linkage of p_s and p_m via (3):

A. *Sex becomes overall cheaper.* Suppose both p_s and p_m drop. If p_m drops, non-sex expenditure x_o must also drop by (1). By the budget constraint, a decrease in p_s, p_m and x_o implies that, for given marital status, a man's consumption of non-reproductive sex must increase – despite the decrease in his marginal utility from non-reproductive sex.

B. *Sex becomes overall more expensive.* Alternatively, male expenditure shifts away from prostitution, raising marriage investment p_m and non-sex consumption x_o. In this case, p_s increases along with p_m – despite the decrease in his marginal utility from non-reproductive sex.

What would be the effect of the Swedish model on trafficking in these cases? Recall that the Swedish model reduces men's (marginal) utility from buying sex from prostitutes and raises the occupational hazards of working as a prostitute (due to the need for secrecy). In case B, an overcompensation effect emerges as in the original model. In case A, the decrease in men's preference for prostitution per se does not create an overcompensation effect. The increase in occupational hazards for prostitutes still does, which can result in an overcompensation effect overall, depending on its relative magnitude.

We will now examine four different examples to see which of the cases emerge. The first example assumes decreasing marginal utility for all three consumption goods. In each of the other three examples, we assume constant marginal utility for one of the goods. In three of the analyzed settings, the Swedish model would lead to an overcompensation effect. In the fourth, cases A and B can coexist as equilibrium outcomes, but only the one corresponding to case B is stable.

Decreasing marginal utilities. We start with an example in which men's marginal utility is decreasing for all consumption goods. Suppose the consumption utilities are $S(x_s) = \ln x_s^\alpha$ for non-reproductive sex, $M(p_m) = \ln p_m^\beta$ for marriage, and $O(x_o) = \ln x_o^\beta$ for the non-sex good. Using the same utility function for non-sex consumption and marriage merely simplifies exposition. We will show that, in an interior equilibrium that satisfies the indifference conditions of married men and women in this setting, decreasing men's marginal utility from non-reproductive sex does *not* reduce the price of non-reproductive sex.

Married men's indifference condition (1) yields

$$\frac{\alpha}{x_s^m} - p_s = \frac{\beta}{p_m} = \frac{\beta}{x_o^m} \tag{5}$$

Conveniently, this implies $p_m = x_o^m$ such that a married man's budget constraint simplifies to

$$y = x_s^m p_s + 2p_m, \tag{6}$$

or $p_m = \frac{1}{2}(y - x_s^m p_s)$. Using this in the women's indifference condition (3) yields

$$\begin{aligned} sp_s &= \frac{1}{2}(y - x_s^m p_s) + w + h \\ p_s &= \frac{\frac{1}{2}y + w + h}{s + \frac{1}{2}x_s^m} \end{aligned} \tag{7}$$

Note that (6), (7), and the first equality in (5) form a system of three equations that yield values for p_s, p_m, and x_s^m in an interior equilibrium. We are interested in how the price of sex p_s varies with α in such an equilibrium.

Now, consider the following conjecture: a decrease in men's marginal utility from non-reproductive sex, i.e., in α, lowers the price of sex p_s. We can disprove this conjecture for an interior equilibrium defined by (5)-(7).

The following must hold for the conjecture to be true: According to (7), the decrease in α must lead to an increase of x_s^m. And by (3), p_m must decrease if p_s decreases. These conditions cannot simultaneously hold. To see this, rearrange the first equality in (5) to $p_s = \frac{\alpha}{x_s^m} - \frac{\beta}{p_m}$ and plug this into (6):

$$y = \alpha - \frac{\beta}{p_m} x_s^m + 2p_m.$$

Suppose α and p_m decrease while x_s^m increases. Then the right-hand side of the above equation strictly decreases, violate the equation, which would imply that married men do not exhaust their budget following the decrease in α. Hence, the conjecture cannot be correct. The alternative conjecture that a decrease in α causes x_s^m to decrease and both p_m and p_s to increase cannot be disproved. Any interior equilibrium in this example must conform to case B, as does the equilibrium in our original model.

Constant marginal utility of non-sex consumption. Now suppose men face a constant marginal utility of non-sex consumption but decreasing marginal utility from marriage and non-reproductive sex. This is, in essence, the baseline setting in Edlund and Korn (2002) who use the non-sex good as the numéraire. In this example, changing men's marginal utility from non-reproductive sex has *no* effect on the price of sex (in an interior equilibrium) due to women's "supply" condition (3), although demand for prostitution decreases.

Suppose men's consumption utilities are $S(x_s) = \ln x_s^\alpha$ for non-reproductive sex, $M(p_m) = \ln p_m^\beta$ for marriage, and $O(x_o) = cx_o$ for non-sex consumption. The men's indifference conditions (1) and (2) yield

$$\frac{\alpha}{x_s^m} - p_s = \frac{\beta}{p_m} = c \Rightarrow p_m = \frac{\beta}{c}, x_s^m = x_s^u = \frac{\alpha}{c + p_s}.$$

Using the above solution for p_m in the women's indifference condition (3) yields

$$p_s = \frac{\frac{\beta}{c} + w + h}{s}.$$

Consider a decrease in men's (marginal) utility from non-reproductive sex, i.e., in α. The prices of marriage and non-reproductive sex do not change. The "tax" on johns merely shifts male spending from non-reproductive sex (x_s^u, x_s^m) to non-sex consumption (x_o^u, x_o^m).

The Swedish model would, in this case, lead to an overcompensation effect. Although p_m is unchanged, the occupational hazards of working as a prostitute increase (i.e., h increases), which in turn increases (the compensating differential Δ in) p_s. Thus the impact of the policy in this setting is qualitatively the same as in our original model.

Constant marginal utility from marital investment. In this example, men's marginal utility is constant for marriage expenditures but decreasing for non-reproductive sex and non-sex consumption. Suppose the consumption utilities are $S(x_s) = \ln x_s^\alpha$ for non-reproductive sex, $M(p_m) = bp_m$ for marriage, and $O(x_o) = \ln x_o^\gamma$ for non-sex consumption. A married man's indifference condition (1) yields

$$\frac{\alpha}{x_s^m} - p_s = b = \frac{\gamma}{x_o^m} \Rightarrow x_o^m = \frac{\gamma}{b}. \tag{8}$$

We conjecture that this setting generates case B since, with non-sex consumption fixed, a shift of married men's spending out of the prostitution market must go fully into the marriage market.

Using (8) to replace x_o^m and (3) to replace p_m, we can write a married man's budget constraint as

$$y = (s + x_s^m)p_s + \frac{\gamma}{b} - w - h. \tag{9}$$

This shows that a decrease in α leads to an increase in p_s if it leads to a decrease in x_s^m. To see that the latter must hold, use the first equality in (8) to replace p_s in (9) and rearrange:

$$\begin{aligned} y &= (s + x_s^m)\left(\frac{\alpha}{x_s^m} - b\right) + \frac{\gamma}{b} - w - h \\ y &= \left(\frac{s}{x_s^m} + 1\right) b\alpha - bx_s^m - w - h + \frac{\gamma}{b} - sb. \end{aligned}$$

The last equality shows that, if α decreases, so must x_s^m for the budget constraint to remain satisfied; married men must shift consumption from non-reproductive sex to marriage. And if x_s^m decreases, (9) implies that p_s increases, which proves the conjecture.

Constant marginal utility from non-reproductive sex. Last, suppose men derive a constant marginal utility from non-reproductive sex and decreasing marginal utility from marriage and non-sex consumption. In this example, there can be two interior equilibria.

Suppose the consumption utilities are $S(x_s) = e$ for non-reproductive sex, $M(p_m) = \ln p_m^\beta$ for marriage, and $O(x_o) = \ln x_o^\gamma$ for non-sex consumption. The married men's indifference condition (1) yields

$$e - p_s = \frac{\beta}{p_m} = \frac{\gamma}{x_o} \Rightarrow p_m = \frac{\beta}{e - p_s}, x_o^m = x_o^u = \frac{\gamma}{e - p_s}. \qquad (10)$$

Using this in the women's indifference condition (3) yields $sp_s - \frac{\beta}{e - p_s} - w - h = 0$, which can be written as the quadratic equation

$$p_s^2 - \frac{se + w + h}{s} p_s + \frac{\beta + (w + h)e}{s} = 0$$

To save on notation, define $k_1 \equiv \frac{se+w+h}{s}$ and $k_2 \equiv \frac{\beta+(w+h)e}{s}$. For all $k_1 > 2\sqrt{k_2}$, this has two (positive) solutions,

$$p_s^* = \frac{k_1}{2} \pm \sqrt{\frac{k_1^2}{4} - k_2},$$

with associated values for the price of marriage and non-sex expenditure,

$$p_m^* = \frac{\beta}{e - p_s^*}, x_o^* = \frac{\gamma}{e - p_s^*},$$

the number of prostitutes, by the market-clearing condition (4),

$$n^* = \frac{y - p_m^* - x_o}{sp_s^* - p_m^*} N$$

and non-sex consumption, by married and unmarried men's budget constraints,

$$x_s^{m*} = \frac{y - p_m^* - x_o}{p_s^*}, x_s^{u*} = \frac{y - x_o}{p_s^*}.$$

Note that, for non-sex consumption x_o, we drop the index for marital status, as $x_o^m = x_o^u$ in equilibrium.

The larger of the two solutions for p_s^* always increases in e and so corresponds to case A. The smaller solution may decrease in e, in which case it corresponds to case B. To see this, consider the parameter values $\beta = .2$, $\gamma = .01$, $y = 1.5$, $s = 1$, and $w + h = 2$. In the table below, we compute the equilibrium values of all endogenous variables for a range of values of e.

		Values of e			
		3.8	3.7	3.6	3.5
Case A	p_s	3.457	3.322	3.174	3.000
	p_m	1.457	1.322	1.174	1.000
	x_o	0.029	0.026	0.024	0.020
	n	0.007	0.076	0.151	0.240
	x_s^m	0.004	0.046	0.095	0.160
	x_s^u	0.426	0.444	0.463	0.493
Case B	p_s	2.343	2.378	2.426	2.500
	p_m	0.343	0.378	0.426	0.500
	x_o	0.007	0.008	0.009	0.010
	n	0.575	0.557	0.533	0.495
	x_s^m	0.491	0.469	0.439	0.396
	x_s^u	0.637	0.628	0.615	0.596

Table 1: Multiple equilibria.

By comparison, in the case-A equilibrium, the prices of marriage and sex are higher, more women are married, and consumption of non-reproductive sex is lower for both married and unmarried men. As a result, the prostitution market is smaller. When men's relative preference for prostitution decreases, the price of sex decreases, but counterintuitively, married and unmarried men increase their consumption of non-reproductive sex, and the number of prostitutes increases. The increased spending in the prostitution market comes at the expense of marital investments and non-sex consumption.

The reverse occurs in the case-B equilibrium. When men's relative preference for prostitution decreases, expenditure shifts away from the prostitution market: both married and unmarried men spend less on non-reproductive sex and more on marriage and non-sex consumption, and more women become married. Hence the prostitution market shrinks. However, the price of sex increases.

Finally, note that only the case-B equilibrium, i.e., the one under the smaller of the two solutions for p_s^*, is stable. To see this, note that individual rationality implies $p_s^* \in [0, e]$; men won't pay more, women won't accept less. Every interior equilibrium hence satisfies $p_s^* \in [0, e]$. Now define

$$\Phi(p_s) \equiv sp_s - \frac{\beta}{e - p_s} - w - h$$

where we use $p_m = \frac{\beta}{e-p_s}$ from a married man's indifference condition (10). $\Phi(p_s)$ reflects women's "career" preference. They strictly prefer prostitution if $\Phi(p_s) > 0$, and marriage if $\Phi(p_s) < 0$. At an interior equilibrium, $\Phi(p_s) = 0$.

Note that $\Phi(0) < 0$ and $\lim_{p_s \to e} \Phi(p_s) < 0$, and recall that $p_s^* \in [0, e]$. This means that $\Phi'(p_s^*) > 0$ for the smaller of the interior solutions and $\Phi'(p_s^*) < 0$ for the larger one. Hence, only the smaller solution identifies a stable equilibrium, which is the case-B equilibrium.

B. Wage heterogeneity

This appendix considers an extension in which the wages women can earn in the "regular" labor market are heterogenous. This changes the supply elasticity of prostitution in a way that creates a countervailing effect to the overcompensation effect highlighted in our baseline model.

Recall that there is a unit mass of women, whom we now index by $i \in [0, 1]$. Let woman i's wage in the regular labor market be given by an increasing "wage function" $w(i)$. We will consider two different specifications of $w(i)$.

Wage heterogeneity as a linear function. Suppose $w(i) = \underline{w} + ai$ where $\underline{w} \geq 0$ is the lowest possible wage and $a \geq 0$ is a slope to capture wage dispersion in a simple way.

Our original model assumes $a = 0$. In that model, an increase in h leads to an increase in p_s (which creates the overcompensation effect for traffickers). We are interested in how a positive a – which makes the supply of prostitution less elastic – modulates the impact of h on p_s. For simplicity, we fix some parameters that are of lesser interest here: $s = 1$, $e/k = 2$, and so $\sigma = se/k = 2$.

Consider a decriminalized market. An interior equilibrium satisfies

$$p_s - h = p_m + w(n) \tag{11}$$

$$\frac{p_s}{e} = \frac{p_m}{k} \tag{12}$$

$$np_s = (1-n)(y - p_m) + ny \tag{13}$$

where the indifference condition (11) identifies the woman who is the *marginal* entrant into the prostitution market, and whose index value equals the number of prostitutes, as all and only women with a lower index value prefer prostitution to marriage.

Indifference conditions (11) and (12) pin down the prices as functions of n:

$$p_s = 2(\underline{\omega} + an) \qquad \text{and} \qquad p_m = \underline{\omega} + an$$

where $\underline{\omega} \equiv \underline{w} + h$. Note, again, that the women's indifference condition forces the prices to move in lockstep. Plugging these prices into market-clearing condition (13) yields the quadratic equation

$$an^2 + (a + \underline{\omega})n + \underline{\omega} - y = 0 \tag{14}$$

An interior equilibrium exists if this equation has (at least) one positive solution $n \in (0, 1)$. A necessary condition for that is a negative intercept:

Assumption 1. $y > \underline{\omega}$.

With a positive intercept, (14) has either no or only negative solutions. With a negative intercept, it has two solutions, one of which is positive and identifies an interior equilibrium if smaller than 1. Suppose the latter solution,

$$n = -\frac{a+\underline{\omega}}{2} + \sqrt{\frac{(a+\underline{\omega})^2}{4} + (y - \underline{\omega})},$$

is a viable equilibrium. The impact of $\underline{\omega}$ on the equilibrium number of prostitutes is

$$\frac{\partial n}{\partial \underline{\omega}} = -\frac{1}{2} + \frac{1}{2} \frac{\frac{(a+\underline{\omega})}{2} - 1}{\sqrt{\frac{(a+\underline{\omega})^2}{4} + (y - \underline{\omega})}} < 0.$$

(The inequality holds under Assumption 1.) That is, an increase in occupational hazards h discourages voluntary prostitution. One can also verify that $\frac{\partial n^2}{\partial \underline{\omega} \partial a} > 0$. Intuitively, an increase in occupational hazards h discourages prostitution less when the outside options of the marginal prostitutes are worse, or more precisely, deteriorate at a higher rate (larger a). In fact, $\lim_{a \to \infty} \frac{\partial n}{\partial \underline{\omega}} = 0$.

With this, note that changes (via h) in $\underline{\omega}$ have two countervailing effects on p_m and p_s. The direct effect is an increase in the price to compensate voluntary prostitutes for the increased hazard. This underlies the overcompensation effect for traffickers in our original model. Under wage heterogeneity, there is further an indirect effect through the term an: since the increase in $\underline{\omega}$ causes exit from prostitution and lowers n, the wage $w(n)$ of the *marginal* prostitute decreases. Formally,

$$\begin{aligned} \frac{dp_m}{d\underline{\omega}} &= \frac{\partial p_m}{\partial \underline{\omega}} + \frac{\partial p_m}{\partial n} \frac{\partial n}{\partial \underline{\omega}} \\ &= 1 + a \left(-\frac{1}{2} + \frac{1}{2} \frac{\frac{a+\underline{\omega}}{2} - 1}{\sqrt{\frac{(a+\underline{\omega})^2}{4} + y - \underline{\omega}}} \right). \end{aligned} \tag{15}$$

While 1 captures the positive effect, the second term is the negative effect. The positive effect prevails when wage dispersion is small: As $a \to 0$, (15) converges to 1. For flatter wage functions, though the change in n is larger, the change in $w(n)$ is small.

This does not necessarily mean that (15) is negative when the wage function is steep. Consider, for example, $\underline{w} = 0$ and $a = y - \underline{w}$. This represents significant dispersion in female (effective) wages, but (15) becomes

$$1 + y \left(-\frac{1}{2} + \frac{1}{2} \frac{\frac{y+h}{2} - 1}{\sqrt{\frac{(y+h)^2}{4} + (y - h)}} \right).$$

Suppose $y = 1$. Then, for $h \to 0$, the above converges to a positive limit:

$$1 + \left(-\frac{1}{2} + \frac{1}{2} \frac{\frac{1}{2} - 1}{\sqrt{\frac{1}{4} + 1}} \right) \approx .276.$$

At the other end, for $h \to y = 1$ ($h > y$ would violate Assumption 1), the limit is positive too:

$$1 + \left(-\frac{1}{2} + \frac{1}{2} \frac{0}{\sqrt{1}} \right) = \frac{1}{2}.$$

Wage heterogeneity as a step function. Suppose there are two groups of women that differ in the wage that they can earn in the regular labor market. A share $\lambda \in (0, 1)$ of the women are face a "low-skill" wage $\underline{w}$, and a share $1 - \lambda$ face a "high-skill" wage $\overline{w} > \underline{w}$. Thus, the wage function is

$$w(i) = \begin{cases} \underline{w} & \text{if } i \leq \lambda \\ \overline{w} & \text{otherwise.} \end{cases}$$

In this example, even if a policy changes the identity of the marginal prostitute, as long as this change occurs within a given population, the overcompensation effect prevails. It is only where the marginal prostitute would switch from being a high-skill to a low-skill worker that a law against prostitution would lower the price of sex. Whether this is likely would depend on parameters, particularly the distribution of education and skill among female workers (λ) and labor market opportunities for women ($\underline{w}, \overline{w}$), which can differ markedly across countries. In practice, when prostitutes have mainly a low-skill background with homogeneous labor market opportunities, we expect the overcompensation effect to prevail.

One can also imagine differences between various types of prostitution (e.g., escorts or streetwalkers), with women working in "high-end" market segments (who tend to be better looking, younger, and healthier) potentially having better outside options. However, *within* those market segments, the opportunity costs of foregone labor income could still be relatively homogeneous. We could capture this situation by assuming that two groups of women with $\underline{w}$ and $\overline{w}$, respectively, differ in the type of non-reproductive sex they can offer, valued by men at $\underline{e}$ and $\overline{e} > \underline{e}$, respectively. This would lead to two prostitution submarkets, one cheaper (more expensive) than the other and served by $\underline{w}$-women ($\overline{w}$-women). In such a model, laws that raise h or reduce $\overline{e}$ and $\underline{e}$ would generate an overcompensation effect in *each* of those two market segments.