

**Monopolistic competition with product  
differentiation in prostitution markets:  
explaining the returns to BMI**

Scott Cunningham

Department of Economics

Baylor University

December 2, 2011

## Abstract

We develop a monopolistic competition theory of prostitution with product differentiation in which male preferences for female appearance and body type determine marriage market assignments and subsequent marital payouts. Prostitution carries different benefits and costs to individual agents that varies according to their attractiveness if male preference for beauty carries into the prostitution markets. We hypothesize that technological shocks facilitating improved coordination between buyer and seller in the sex markets will select agents into prostitution who have superior outside marital options, and to protect those options, they will typically sort into practices and service offerings that manage that overall risk. We test our theory using a variety of datasets and estimation strategies. We find that technological shocks lowering the risk of detection for sex market participants caused an increase in more attractive prostitutes, and potentially clients as well. Using a novel survey conducted of internet-based prostitutes, we estimate the differential earnings by body type and the service offerings. Underweight (overweight) prostitutes earned 84% more (70%) less in a week than comparable prostitutes with normal BMI due to different higher (lower) weekly client visits, different prices of individual services, differences in session length and differences in client types. Women with ideal body type charged higher prices to older males. Obese prostitutes charged 22 percent less for anal sex with a condom, and 59 percent more for anal sex without a condom. Law enforcement and health researchers would benefit from a more careful understanding of how technology has reallocated STI risk and victimization across the underground market for commercial sex.

# 1 Introduction

Beauty plays a well-known role in the determination of earnings, schooling, marriage and even sexual choices, but little is known about its importance in prostitution. Theoretically, the effect is complicated and nuanced. Prostitution carries a penalty in the marriage market, and thus more attractive males and females face higher costs of entry (Edlund and Korn, 2002; Guista, Tommaso and Strom, 2009). But, if males prefer women who are more attractive, then returns to appearance raise the marginal product of attractive prostitutes (Gertler, Shah and Bertozzi, 2005). Understanding of male preferences for beauty helps account for the systematic feedback between marriage markets and prostitution markets, which in turn helps us understand the role of BMI in the determination of risky sex among prostitutes, as well as the impact that technology such as broadband and various internet service providers has had on the market structure for commercial sex. We develop a monopolistic competition theory of prostitution to model the feedback between marriage and prostitution via individual appearance and body type to study the market structure for commercial sex markets and to understand the incentives to engage in high risk sexual behaviors. We empirically test our theory using a variety of datasets and estimation strategies.

We estimate the reduced form semi-elasticity of weekly earnings with respect to body mass index (BMI) score is -0.05. Women with an underweight BMI classification earned 84 percent more than women with normal BMI; women with an overweight BMI earned 70 percent less than women with a normal BMI. We show that the inelastic BMI penalty is primarily driven by differential sorting by prostitute body type into different product markets best explained with vertical and horizontal differentiation. The difference in the markets is primarily seen in client characteristics and transaction details. Women with normal BMI operate in higher quality product markets characterized by price discrimination through bundling, careful management of risk, and fewer clients. Women in this market typically price discriminate using the age of the client as the hurdle, and charge older males higher prices, with the highest price given to 55 year olds. We interpret this as evidence that older males have more income, and are likely the marginal entrants into

prostitution at the extensive margins created by technology-driven falling probabilities of detection. The higher quality products are services that typically last considerably longer than the lower quality products, which may suggest that more attractive prostitutes use bundling/tie-in strategies to price discriminate as well. We also find evidence for this group being marginally more sensitive to the risks associated with new clients - both in that they saw fewer clients overall, including 0.45 fewer “new” clients compared to underweight prostitutes, and more likely than underweight and obese prostitutes to perform a “background check” on the new clients. Regular clients of a normal BMI prostitute paid 4.7% lower prices, as well, perhaps suggesting that information about client type is valuable in these markets.

Underweight, overweight and obese prostitutes appear to operate in more competitive “lower quality” product markets and face one another primarily as rivals. All three groups spent between 27 and 38 percent less time overall with their clients than women with normal BMI. Perhaps because underweight prostitutes are closer to the ideal, relatively speaking, than either overweight or obese, male preferences for more attractive prostitutes lead underweight women to capture most of the market. Underweight prostitutes saw 1.08 more regular clients and 0.55 more new clients per week than women with normal BMI, whereas obese prostitutes saw 0.45 fewer regulars than normal BMI. Overweight and obese prostitutes must compete on either price, therefore, or other margins to capture the marginal clients, and we find that showing up primarily by taking more chances with new clients, lowering prices on infra-marginal services, such as anal sex with a condom, and providing higher risk services at a compensating differential like anal sex without a condom. We estimate 22 percent lower payments for anal sex with a condom and a 59 percent higher price for anal sex without a condom for obese prostitutes using a fixed effects modeling strategy.

Section two explains the relevant background for our study; section three is our theory of monopolistic competition; section four describes the data; section five presents our empirical results; section six is the conclusion.

## 2 Background

A number of studies have shown that appearance and body type affects educational attainment (Sabia, 2007), labor market earnings (Hamermesh and Biddle, 1994; Cawley, 2004; Fletcher, 2009), marriage outcomes (Averett and Korenman, 1996; Mukhopadhyay, 2008), and even selection into crime (Mocan and Tekin, 2010). The impact of BMI on females is particularly robust across numerous studies, methodologies, and datasets (Sabia and Rees, 2011). A more recent literature has shown differences in inframarginal sexual choices by overweight adolescent females. Overweight adolescent females delay vaginal sex (Cawley, Joyner and Sobal, 2006; Sabia and Rees, 2011), but conditional on engaging in any sexual behavior are more likely to reportedly engage in anal sex (Averett, Corman and Reichman, 2012).

Insofar as BMI and appearance have associated causal returns and penalties, then non-idealized body type may have have an effect on prostitution markets as well. The insight that prostitution and marriage markets are connected in non-complementary ways was first suggested by Edlund and Korn (2002). The authors' explanation for high wages in prostitution despite low skill, labor intensive, female-concentrated work is that husband demand for paternity makes prostitutes poor candidates for wives. Therefore, prostitution wages must compensate the marginal prostitute her opportunity costs, which are both the costs of the work itself as well as the foregone marital surplus. Evidence for Edlund and Korn (2002)'s marriage explanation of prostitution has been mixed. Shah and Arunachalam (2008) showed that prostitutes were more likely to be married than the counterfactual non-prostitute female in Latin America. Cunningham and Kendall (2011*a*), on the other hand, examined Edlund and Korn (2002)'s prediction that a temporary increase in non-residential males ("men in transit") should cause an increase in prostitution relative to marriage using the 2008 Democratic and Republican National Conventions as a natural experiment. For each city, the convention caused an increase in total prostitution advertising at Craigslist's now-defunct "erotic services" relative to two control group cities.

A separate though not mutually exclusive theory of prostitution is Guista, Tommaso

and Strom (2009). The authors incorporate endogenous stigma costs into client demand and supply assuming homogenous products. As the size of the market for commercial sex grows, the reputational penalty received from prostitution exchange declines, and in turn causes the market to grow even more before reaching the new equilibrium. Given the endogeneity of stigma to the size of the market, exogenous factors lowering the probability of detection might have larger feedback effects on the market's size than might be expected. Some evidence exists that technology functions in this way. For instance, Cunningham and Kendall (2011*b*) estimated that a negative effect of expanding online prostitution markets on street prostitution arrests using broadband dispersion as an instrumental variables with the effects strongest for female arrestees aged 35-44. Logan and Shah (2009) model the use of information technology by gay male escorts as a signaling device to facilitate coordination in the underground market.<sup>1</sup>

One common theme, though, is that it's historically believed to have been linked to the spread of STDs, as well as poverty more generally. This is partly why the most important empirical studies of prostitution have focused on poverty and unprotected sex in an effort to test for the existence of compensating differentials for higher risk sex. Rao, Gupta, Lokshin and Jana (2003) estimated that sex workers practicing safe sex consistently earned 70% less than those women who did not. Similarly, Gertler, Shah and Bertozzi (2005) estimated a 23% wage premium for unprotected sex among Mexican prostitutes on average, but a 50% premium if the prostitute was viewed as attractive. Robinson and Yeh (2011*a,b*) show that the decision to supply risky sex is typically a dynamic decision triggered by negative income shocks, and may be a consequence of poor functioning credit markets creating the need for consumption smoothing through other channels.

---

<sup>1</sup>Gambetta (2010) analyzes extensively the many ingenious methods criminals rely upon to coordinate, including signaling, credible threats, and reputations. Prostitution is not covered, though, in the book.

### 3 Monopolistic Competition Theory

Our model of prostitution is a partial equilibrium in which male preferences for idealized body types have direct and indirect effects on commercial sex market structure. We first model the marital assignment using a simple two-sided matching model in which males and females are assigned to marriage through the deferred acceptance algorithm, and receive utility from marriage based on their spouse’s location in their preference ordering. Assuming non-prostitutes/non-clients are strictly preferred to prostitutes/clients, marital payouts, and therefore marital stigma costs, will be higher for males and females with more attractive body types. The rest of the model is a simple extrapolation of Salop (1979) with and without vertical differentiation to explain the sorting of men and women by body type and appearance into prostitution can be usefully used to explain the differentiation of risky behaviors across the market place.

#### 3.1 Preferences and the marriage assignment

We describe a “marriage market” as a two-sided matching problem of equal number males to females where each male (female) rank orders each female (male) according to a partially correlated strict and complete preference function that satisfies two rules (Roth and Sotomayor, 1990). First, males (females) strictly prefer non-prostitutes (non-johns) to prostitutes (johns):<sup>2</sup>

$$\begin{aligned} \text{Females: } U(NJ) &> U(J) \\ \text{Males: } U(NP) &> U(P) \end{aligned} \tag{1}$$

Second, males (females) strictly prefer females (males) whose body type is closer to a broadly defined ideal body type,  $t^*$ . We model a person’s body type,  $t$ , as distance from

---

<sup>2</sup>Equation 1 uses “P” and “NP” to identify prostitute and non-prostitute females, respectively, and “J” and “NJ” to identify male prostitute consumers and non-consumers, respectively. We “J” for male customers because traditionally, prostitutes would call their clients “johns”. Prostitutes are always females in our model, and johns always males. See Logan (2010) for a detailed examination of male escort markets.

the ideal, or  $|t - t^*|$ .<sup>3</sup>

$$\begin{aligned} \text{Females: } U(t^*|NJ) &> U(|t - t^*||NJ) > U(t^*|J) > U(|t - t^*||J) \\ \text{Males: } U(t^*|NP) &> U(|t - t^*||NP) > U(t^*|P) > U(|t - t^*||P) \end{aligned} \quad (2)$$

Males and females are assigned to marriages based on their own preferences, everyone else's preferences, and the "deferred acceptance algorithm" (Roth, 2008). Each round begins with each male making a marriage proposal to the female highest in his preference ordering who has not previously rejected his offer. Each female who receives an offer in a round accepts it if it exceeds her reservation utility, otherwise she rejects it. If she has two acceptable offers, she accepts the more preferred offer and rejects the other.<sup>4</sup> Any male whose offer was rejected waits until the subsequent round wherein he then offers to his second preferred. The process is repeated until every male runs out of women to whom to propose or until their offers are accepted, resulting in the marital assignment  $\mu_m$ .

In their seminal study, Gale and Shapley (1962) showed that if players in a two-sided marriage market had strict, complete and continuous preferences, then the deferred acceptance algorithm's marital assignment,  $\mu_m$ , was non-empty and stable such that there did not exist any two people in  $\mu_m$  who preferred one another to their assigned marriage.<sup>5</sup> As equation 2 creates positively correlated preferences for ideal body type and appearance, the marriages in  $\mu_m$  will positively sort on body types that are closer to the ideal. We define the marital payouts from marriage as  $V(\mu_m)$ , which given the positive sorting on body type and non-prostitution status delivered by the deferred acceptance algorithm,  $V$  will be higher for non-prostitutes than prostitutes (equation 1) and higher for more attractive individuals (equation 2).

---

<sup>3</sup>We proxy for idealized body type with a body mass index (BMI) range of 18.5 to 25 ("normal").

<sup>4</sup>Marriage is, in other words, one-to-one in the marriage market.

<sup>5</sup>The one caveat is that the stability outcome depends on whether men ask women for marriage ("m-optimal") or women ask men ("f-optimal"). The f-optimal marital assignment is that  $\mu_f$  assignment wherein women had proposed at each round, and men had accepted and rejected offers. Roth and Sotomayor (1990) shows that the f-optimal assignment,  $\mu_f$  is weakly preferred by the elements of F to the M-optimal assignment,  $\mu_m$ , and vice versa for males. Which is to say, whereas females may have the same ranked husband in either  $\mu_m$  or  $\mu_f$ , no female will have a lower ranked husband in  $\mu_f$ .



### 3.2 Marital stigma

If a non-prostitute or non-john engages in prostitution exchange, and is discovered, then that person is reassigned to a new marriage satisfying equations 1- 2, leading in equilibrium to the new, stable assignment,  $\widehat{\mu}_m$  and corresponding marital payout  $V(\widehat{\mu}_m)$ .<sup>6</sup> Given equations 1-2, then marital payouts upon discovery will satisfy the following conditions:

$$V(\mu_m) > V(\widehat{\mu}_m) \tag{3}$$

$$V[(\mu_m - \widehat{\mu}_m)|t^*] \geq V[(\mu_m - \widehat{\mu}_m)||t - t^*|] \tag{4}$$

Equation 3 is according to the allocations created by the deferred acceptance algorithm. If  $\mu_m$  was the best allocation for a person given the preferences of everyone in the population, and detection causes a reranking according to equations 1-2, then equation 3 holds for most people. Each non-prostitute, regardless of body type, is made strictly worse off from discovery. Equation 4 states that the loss in marital payouts for women with body types closer to the ideal will be weakly larger than the loss to those women far from the ideal. We can rearrange equation 3 to represent the stigma costs, or “marriage market penalty”, from prostitution as  $\mathbf{A} = V(\mu_m) - V(\widehat{\mu}_m)$ .<sup>7</sup> The presence of marital stigmas in prostitution does not mean prostitutes (johns) will never marry. It means that prostitute/john marriages have lower utility. This may help explain Shah and Arunachalam (2008)’s empirical findings regarding a high rate of marriage among prostitutes.

### 3.3 Circular prostitution markets

We model prostitution using Salop (1979)’s circular city because it is a useful benchmark for industries with nonuniform preferences across an empirically validated product space. We assume that there are  $n$  different prostitutes, or “brands”, of a differentiated commodity,  $q$ , available at prices  $p_t$  and locations  $t$  around a circle of unit-circumference.

---

<sup>6</sup>We ignore the externalities that discovery would impose on the other members of the market for simplicity.

<sup>7</sup>To help the reader with our notation, we symbolize the marriage market penalty from prostitution with an  $A$  as a visual reminder of the scarlet letter “A” Hester Prynne was forced to wear publicly as punishment for adultery Hawthorne (1850).

Consumers (“johns”) purchase one or zero unit of the differentiated commodity according to preferences, price and the distribution of brands in the product space. The market for commercial sex in this model is divided into two sectors: a monopoly sector where individual prostitutes do not compete with one another due to high transportation costs and a monopolistically competitive industry with product differentiation.

### 3.3.1 John behavior

Johns are drawn from the set of males in  $\mu_m$ . While they may not be married, they have the potential for a marital match equal to  $\mu_m$ . If they decide to purchase prostitution services,  $q$ , they do so only if they can be compensated their costs, which include any disruptions in the marriage market that could occur if detected. Formally, a john purchases  $q$  and receives utility from sex,  $u$ , minus the brand’s distance absolute distance from his ideal body type at a rate of  $c$  per unit traveled:

$$U(t, t^*) = u - c|t - t^*| \tag{5}$$

where  $|t - t^*|$  is the shortest arc length between the two points. The john chooses to maximize consumer surplus choosing optimal brand,  $t$ , net of the price,  $p_t$ , and the expected marriage penalty associated with that prostitute:

$$\max_t U_j(t, t_j^*) - p_t - \psi_t A_j \geq \bar{u} \tag{6}$$

where  $\psi_t$  is the probability of detection bounded between 0 and 1 associated with the prostitute,  $A_j$  is the marital penalty for that john if detected, and  $\bar{u}$  is the maximum consumer surplus from his next best alternative activity.

Assume that there exists a price,  $p^*$ , for a prostitute located at  $t = t^*$  that would maximize consumer surplus. This price is the john’s reservation price,  $v$ , equalling

$$v = u - \bar{u} - \psi_t A_j \tag{7}$$

Equation 7 shows that the john’s willingness to pay for services is an increasing function of

the gross utility from service,  $u$ , and decreasing in  $\bar{u}$ ,  $\psi_t$  and  $A_j$ . Substituting equations 5 and 7 into 6, yields

$$\max_t v - c|t - t^*| - p_t \geq 0 \quad (8)$$

Restated, the john's problem is to select service from a local monopolist if and only if consumer surplus associated with that particular brand is non-negative. Otherwise, the john simply chooses to consume the outside homogenous good,  $\bar{u}$ .

### 3.3.2 Monopoly behavior

We model prostitution markets using Salop (1979)'s circle model. We begin with a model of monopoly. Prostitutes females from the set of matches in  $\mu_m$ , and therefore have a marital stigma,  $A_t$ , which is higher for women whose body type are closer to the ideal,  $t^*$ . Prostitutes locate around a circle of unit circumference and prefer to locate far from one another so as to enjoy market power. Insofar as a john has no preferences for body type, then the john can travel anywhere in the city at zero cost ( $c = 0$ ). But, preferences over beauty and body type translate into positive transportation costs. First we consider the demand for a homogenous service.

Consider a john located at  $x = |t - t^*|$  where  $t$  is a prostitute selling a service,  $q$ , valued at  $u$  at price  $p_t$ . The john is willing to buy from the prostitute at his location,  $x$ , if and only if  $v_t - cx - p_t \geq 0$  (equation 8). Using this expression, the maximum distance,  $x_m$ , he will travel to purchase from a prostitute is:

$$x_m = \frac{v_t - p_t}{c} \quad (9)$$

If there are  $J$  johns around the circle, then in the absence of competition, the prostitute captures all johns  $x_m$  distance in either direction ( $2Jx_m$ ). Substituting equation 9 into  $2Jx_m$  yields her potential market demand

$$q_m = \frac{2J}{c}(v_t - p_t) \quad (10)$$

The prostitute supplies  $q$  at cost,  $C_t(q)$  composed of fixed and variable costs: an expected

cost per client of  $\psi_t A_t$ , a certain cost per client of  $m$ , and a fixed cost of  $F$ :

$$C_t(q) = F + (\psi_t A_t + m)q \quad (11)$$

Using equation 7 and 11, and ignoring fixed costs, we see that the market participants will be defined by those individuals for whom the weighted marriage penalties is less than total gains from trade:

$$\psi_t(A_j + A_t) \leq u - \bar{u} - m \quad (12)$$

Equation 12 helps us understand the composition of buyers and sellers under different probabilities of detection. Only johns for whom  $v - cx \geq p_m^*$  will purchase from the vendor at  $p_m^*$ . We would expect that when the probability of detection is high,  $u$  must be large and/or sellers and buyers are located sufficiently far from ideal points. When the probability of detection is falling, though, then markets may consist of buyers and sellers located closer to ideal points.

If  $q_h$  and  $q_l$  differ with regards to the distance a john will travel, then we would expect local monopolies to form more often around those prostitutes with higher  $A_t$ . The maximum distance a john would travel for the two types of goods would be

$$x_h = \frac{v - p}{c_h} \quad (13)$$

$$x_l = \frac{v - p}{c_l} \quad (14)$$

If  $c_h > c_l$ , then necessarily  $x_h$  is a shorter arc distance from  $t^*$  than  $x_l$ . Assuming the johns purchase one unit of either  $q_h$  or  $q_l$ , then the monopolist facing  $q_h = \frac{2J}{c_h}(v - p)$  would service this market, leaving residual demand to be provided by remaining prostitutes. We would expect therefore for monopolist pricing, price discrimination, bundling and other practices in the  $x_h$  range of the prostitution market given the higher likelihood that they are able to operate as monopolists.

Given the higher marital stigma that prostitutes with higher values of  $A_t$  incur as a result of their closer proximity to  $t^*$ , we would expect prices to be higher for those more attractive women and/or prostitutes. Thus, we would expect that only those males who

can pay for the higher monopoly prices at  $x_h$  will do so, which may suggest that we would expect to see a compositional difference in the males at that margin. Therefore, higher prices in the monopoly region would select upon men of higher income, potentially. The residual demand markets would therefore consist of johns with a lower willingness to pay.

### 3.3.3 Competitive equilibrium

In the previous analysis, we assumed that a prostitute was located on the circle such that the transportation costs to the next nearest prostitute was too high, creating a local monopoly on that arc of the circle. An increase in the number of prostitutes would consequently require less distance between any two, though, and create the potential for two monopoly markets to overlap for a john who is willing to travel to either vendor. Continuing to assume unit-circumference and  $n$  firms, then each prostitute will locate  $\frac{1}{n}$  distance from each other on the circle.<sup>8</sup> Assume that the brand's closest competitor is  $\frac{1}{n}$  away and charges  $\underline{p}$ . If the prostitute charged  $p$ , then she sells to all johns located within a distance  $x_c$ , where  $x_c$  is defined as the distance where her market just reaches her rivals' market, and at  $x_c$ , a john is indifferent between purchasing from either prostitute.<sup>9</sup> This indifference condition requires the john receive the same net utility from either vendor, or:

$$v - cx_c - p = v - c\left(\frac{1}{n} - x_c\right) - \underline{p} \quad (15)$$

where the left-hand-side of equation 15 is the net utility the john receives from purchasing at  $x_c$  at  $p$  and the right-hand-side is the net utility the same john receives from purchasing at  $\frac{1}{n} - x_c$  at  $\underline{p}$ . Note that this implies that without the rival brand at  $\frac{1}{n}$ , the monopolist's entire market would be  $x_m$ , and  $x_m > x_c$  because  $x_c$  is that point where the two rivals' market share overlap. Solving equation 15 for  $x_c$  yields a competitive demand equation,  $q_c$ , equalling  $2x_cJ$ , or:

$$q_c = \frac{J}{c} \left( \frac{c}{n} + \underline{p} - p \right) \quad (16)$$

---

<sup>8</sup>While unrealistic, we assume that relocation by firms is costless.

<sup>9</sup>The subscript  $c$  refers to the competitive equilibrium.

Equation 10 and 16 are the two demand equations associated with the brand over two separate ranges. At higher prices,  $p_m$ , the prostitute faces equation 10 and charges monopoly prices. At prices below  $p_m$ , though, equation 16 represents the demand equation. Note that the price of the rival firm enters directly into the decision of the representative prostitute.

Assume that there exists two prostitutes whose body type differs with regards to its distance from  $t^*$ : an obese and non-obese prostitute. The obese prostitute is located a further distance from  $t^*$ . There exists a price,  $p_o$ , such that this john is indifferent between purchasing from the obese prostitute, which holding constant the subjective utility from services,  $u$ , and transportation costs,  $c$ , requires that  $p_o < \underline{p}$ , where  $\underline{p}$  is the price of the rival firm closer to the john's preference. Consequently, for a given service,  $u$ , johns will only purchase services from prostitutes with body types further from the ideal if there exists overlapping markets with rivals and  $p_o < p_c$ .

Sellers in the long run enter into the market only if the total revenue from production exceeds their total costs. Given correlated male and female preferences for ideal body-type, this implies that when  $\psi_t$  is approaching implies that the market for commercial sex if it exists at all will consist of those individuals for whom  $A_j$  and  $A_t$  are sufficiently small under high values of  $\psi_t$ . That is, if the probability of detection is approximately 1 - for instance, because assignation occurs in publicly locations such as streets and highways where visibility and therefore detection is high - then only those who do not stand to lose much in the marriage market will enter, which in our model would imply a greater proportion of individuals drawn from the population of males and females who are the furthest distance from the ideal body type. Consequently, those males most sensitive to the body type ideal will not purchase services, because equation 12 will not hold.

## 4 Description of data

### 4.1 The Erotic Review

The Erotic Review (TER) is a national prostitution review board used by clients since 1998 to share information about prostitutes across the United States and numerous international cities. Similar to Yelp.com, TER uses online reputational mechanisms to increase awareness of quality differences across vendors. TER shares information about prostitutes, such as their characteristics, with members and stores the information online to help others search for prostitutes whose attributes and characteristics were a good fit for the client. The website has operated legally and openly since 1998, and we acquired the data using a PERL script in 2008 to harvest over 500,000 reviews of more than 94,000 prostitutes. A more detailed description of the data is included in the appendix.

### 4.2 Survey for Adult Service Providers

Though TER contains rich information on physical characteristics, it lacks many important variables that we could use to better understand the product space of the commercial sex market necessary to test our model’s predictions. To complement it, we collected data on prostitutes using a field survey entitled the Survey for Adult Service Providers (SASP). While we describe the data in greater detail in the appendix, before explaining our empirical strategy for testing a variety of our predictions, we briefly review some relevant details of the dataset.<sup>10</sup>

Respondents were asked about two dimensions of their recent work experience: transaction-rich information about up to the last 5 client assignments (“client file”) and non-client-specific information related to earnings, labor market experience and background charac-

---

<sup>10</sup>The non-terminated email addresses in our TER served as an approximation of the population ( $n=13,333$ ). From 2008-2009, we contacted respondents by email and telephone requesting their participation in the study, which yielded 685 respondents and a 5.14% response rate, which we argue is the lower bound response rate given we cannot identify whether all emails were operational, only whether they “bounced back”. First we drop all male respondents (14 observations), non-escorts (30 observations), and women whose base of operations are in Canada (40 observations), the United Kingdom (7 observations), or another foreign country (7 observations). These selective criteria left us with 587 observations. Some respondents did not answer all of the questionnaire, and therefore of the 587, we lost another 55 due to missing data.

teristics (“provider file”). As the client file contains repeated observations for the same prostitute, we are able to estimate fixed effects models to control for unobserved heterogeneity.

In the provider file, respondents were asked to self-report height and weight from which we calculated each respondent’s body mass index (BMI) as  $\frac{\text{Weight in pounds} \times 703}{\text{Height}^2}$ . Most of our analysis will focus on four clinical BMI categories: underweight (BMI less than 18.5), normal (BMI between 18.5 and 25), overweight (BMI between 25 and 30), and obese (BMI greater than 30). Our analysis suggests that “normal BMI” is closer to  $t^*$ , underweight is next closest, and overweight/obese is the furthest distance. We report the summary statistics from the provider file using our regression sample only in Table 1. The mean BMI is 23.3, with 62 percent of respondents in the normal range, and 24 percent either overweight or obese. Average age is 28 and years since first entry into prostitution is 5.4, though not all respondents worked continuously since first entry. Our sample is also well-educated – 39 percent had a college degree and another 43 percent had at least some college.<sup>11</sup> Thirteen percent were married and living with their spouses at the time of the survey, and 19.2 percent were cohabiting.

Respondents reported relatively high weekly earnings from prostitution. While 26 percent of the sample claimed they did not see a client the previous week, of the 74 percent who had, the average earnings was \$2,536 (in untaxed income).<sup>12</sup> If we include even those who did not work, the weekly earnings were \$1,846. Including zero counts, respondents saw on average 5.7 clients total in a week, 2.98 of whom were “regulars”.

Respondents were asked to rank their concern or beliefs about the risks of detection, HIV infection and client-initiated assault on a scale of 1 to 10. Concerns about police detection, family detection and client violent were equally weighted around 4, but beliefs about HIV infection were lower by over 1 point (2.7). Fourteen to 15 percent of the sample said they had been assaulted by a client or solicited from a street location, respectively, and 79 percent said they took precautions before meeting with a new client. The kinds of precautions reported were using a third party as a reference (60 percent), using search

---

<sup>11</sup>Signs of formal schooling also show up in parental education.

<sup>12</sup>Gift transfers were not uncommon, but were not coded. Thus weekly earnings consists only of money income, not gifts.



engines and the internet to learn about a new client beforehand (57 percent), requiring clients to show state ID (32 percent) and conducting a background check (21 percent).

We report the summary statistics of the panel dimension of our client file in table 2. Respondents were asked the total payment received for a session, the length of the session in minutes, from which we calculated the hourly adjusted price of a transaction. Respondents are spending on average 2 hours with clients, and receiving \$346 in hourly compensation. The most common session is 60 minutes long (38.1 percent), while 28 percent are shorter than an hour, and 34 percent last longer. Using the entire sample of transactions provided to us, we calculated that 54 percent were sessions with a “regular” client, which is defined as a client the provider had met before or considered one of her “regulars”. Most ( 81.1 percent) of the transactions we analyze are with white clients; black clients appear in only 5.4 percent of all cases, Asian clients in 7 percent, and all other races 6.5 percent. In 5.7 percent of all transactions, the respondent reportedly worked with at least one other provider in the session. Kissing is common in these transactions (55-59 percent of all transactions), and 46.5 percent of the sessions occurred at a hotel room. Traveling out of state for assignation occurs in 16.5 percent of all cases. The types of services provided that we analyze range from massages to penetrative sexual intercourse. Those mean values are listed in the table, as well. As can be seen, unprotected sex of any kind is infrequent, except for unprotected fellatio which occurs in half of all recorded sessions. Ordinarily, vaginal sex is provided, and primarily with a condom worn. Anal intercourse with a condom and without a condom are both rare (5 and 1 percent, respectively).

### **4.3 General Social Survey**

While SASP is a historically unique convenience sampling of the off-street prostitution markets in an industrialized country, it is nonetheless a convenience sample. We attempt to correct for self-selection in particular by estimating the probability of having ever worked as a prostitute using a probit model with a nationally representative dataset and impute the estimated inverse mills ratio to respondents in SASP using those exogenous

covariates shared with the nationally representative dataset. The only nationally representative dataset containing detailed information on respondents in the US over the time period in questions is the General Social Survey (GSS), which is a repeated semi-annual cross-section consisting of between 1,200 and 2,500 respondents semi-annually. We pooled data from the 1991-2010 waves and dropped all male respondents and missing observations yielding a sample of 12,886 females from 1991–2010.<sup>13</sup> Table 3 describes the data used from this analysis. Only 1.83 percent of the female respondents sampled over the time period stated affirmatively they had ever received compensation for sex. The average of the sample is 45 and modal educational attainment for the period is high school education. Consistent with the distribution of race in the US, 79.4 percent of females sampled are white, and 14.2 percent are black. Most of the sample is married (49 percent). Educational attainment among mother and father show that a minority of the sample come from families with a college degree (11 and 13 percent have mothers and fathers with at least a college degree, respectively).

## 5 Estimation strategies and results

We present our findings separately according to the hypothesis of interest. First, we report correlations between prostitution and background characteristics using the GSS sample, from which we estimated the inverse mills ratio for other models. We also report correlations between BMI and prostitution outcomes, such as earning, using the SASP. Secondly, we report our findings of the effect of changing technology present both correlation output between BMI and various associated background characteristics as well as illicit outcomes, and results that are more explicitly test of our model’s hypotheses so that our overall results can be better understood.

---

<sup>13</sup>The GSS asked two questions regarding respondents’ experiences with prostitution. In addition to the one we use, respondents were asked whether they had sex for pay last year. Only 10 respondents of the 12,886 in our GSS sample answered in the affirmative. In comparison, there were 253 female respondents out of 12,886 who said they had *ever* been paid for sex since becoming 18. Waves 1991, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008 and 2010 contained the question regarding prostitution entry.

## 5.1 Selection and inverse mills ratio

The SASP data file is a non-random convenience sample of prostitutes who have reviews at TER. To address the selection bias, we took two strategies. First, we created sampling weights based on the age and race shares of the TER sample and the SASP to create an inverse probability weight that the respondent would have appeared in our data. All regression models and summary statistics are population estimates using the inverse probability weights for analysis. Second, we attempt to correct for self-selection into prostitution based on observable covariates that the GSS and SASP datasets have in common. Unfortunately, GSS does not have information on body type.

First we estimated the probability of selection into prostitution using the GSS controlling for covariates, such as race, age, education, parental education and residential fixed effects using census region, with the following probit model:

$$\text{Prob}(Y_{dt} = 1|\mathbf{x}_{dt}) = \Phi(\beta\mathbf{x}_{dt} + \delta_t + \epsilon_{dt}) \quad (17)$$

where  $\delta_t$  are year dummies. For convenient interpretation, we report the marginal effects in Table 7. The expected probability of having ever worked as a prostitute is concave in age.<sup>14</sup> Black females were 2.8 percent more likely than whites to have said they worked as a prostitute. Higher education attainment lowered the probability of entry by 0.7 percent for some college and 0.6 percent for those with a college degree.<sup>15</sup> Maternal education also lowered the probability of education, though nonlinearly. Compared to mothers with a high school diploma, those whose mothers did not complete high school were 0.6 percent less likely to have entered prostitution, while those whose mother had only some college were 0.9 percent less likely.

We calculated the inverse mills ratio using the fitted values,  $\widehat{\beta\mathbf{X}}$ , from equation 17 as  $\frac{\phi(\widehat{\beta\mathbf{X}})}{\Phi(\widehat{\beta\mathbf{X}})}$ , which is the ratio of the standard normal probability distribution function to the cumulative distribution for the fitted values. We then assigned each respondent in SASP the inverse mills ratios estimated from the GSS sample using the covariates from our

---

<sup>14</sup>The estimated coefficients on age and age-squared are 0.0013391 and  $-0.0000149$ . Using these coefficients, the max of the age parabola is  $\frac{-\beta_{age}}{2 \times \beta_{age^2}}$ , or 45.37 years of age.

<sup>15</sup>See Cunningham and Kendall (2010) for more detailed analysis of education and prostitution.

selection model to identify the nearest neighbor in a linear regression model. All models estimated using SASP control for selection using this imputed inverse mills ratio. Because prostitution street markets declined from the mid-1990s to the present (Cunningham and Kendall, 2011*b*), the year dummies are the excludable variables used to estimate the inverse mills ratio.

## 5.2 BMI and Earnings

Next we analyzed the background characteristics of women with different BMI levels using the following probit model:

$$\text{Prob}(Y = BMI_i) = \alpha + \gamma \mathbf{x}_i + u_i \quad (18)$$

where  $i$  categorizes a prostitute’s weight class as “underweight”, “normal”, “overweight” or “obese” based on her self-reported weight and height used to calculate BMI. Each model was estimated four separate times. All coefficients are presented as marginal effects and reported in Table 8.<sup>16</sup>

Underweight prostitutes in our sample were typically uneducated, less likely to be currently married, and less likely to be cohabiting. They were 52 percent more likely to have a father whose highest level of education was less than a high school degree. Women with a normal BMI constitute the majority of all cases in our sample. If a woman is widowed, she is 33 percent more likely to have a normal BMI. And if her father had less than a high school degree, she was 42 percent less likely to have normal BMI. Overweight prostitutes were 13 percent less likely to be black, and 20 percent more likely to have had at least some college schooling. If married, she was 15 percent more likely to be in the overweight portion of the BMI distribution. And if her father had some college schooling or a college degree, she was 21 to 11 percent more likely to be in the overweight range of BMI. Our estimate of the inverse mills ratio is negative and marginally significant for overweight prostitutes, suggesting that the marginally selected prostitute is less likely to be overweight at the time of the survey. Obese prostitutes follow a unique age-probability

---

<sup>16</sup>We will usually not discuss statistically insignificant for the sake of brevity.

profile wherein the conditional probability of being obese is increasing in age up to age 42, beyond which the conditional probability falls.

Next, we estimated the mean return to BMI<sub>*i*</sub> using the following model estimated with OLS:

$$\ln(\text{earnings}_i) = \alpha + \beta_1(\text{BMI}_i) + \beta_2\mathbf{x}_i + u_i \quad (19)$$

Our results are presented in table 9.<sup>17</sup> A 1-unit increase in respondent BMI was associated with declining weekly earnings of 5 percent. Examining the association using categorical measures of the BMI distribution, we find this is driven primarily by the differences in the two tails of the BMI distribution, underweight and obesity. Underweight prostitutes earned 84 percent more than their counterparts with normal BMI which, using the \$2,536 in mean weekly earnings reported in table 1, is \$4,667 a week. Obese prostitutes, on the other hand, earned 70 percent less than their normal BMI counterparts which using the same mean value of earnings amounts to approximately \$761 per week.

### 5.3 Hypothesis 1: Technology, marriage penalties and compositional effects

Our first hypothesis is deduced from equation 12 and 13. Women and men who are more attractive have higher values of  $A$  due to appearance's return in the marriage and labor markets. High expected damages from discovery keep attractive women (men) from entering, or conditional on entry, cause attractive women to sort into  $q_h$  markets where they face greater market power assuming such markets can be sustained. Because  $x_h < x_l$ , women operating in this market will see fewer clients. If males with higher values of  $A_j$  have higher incomes, then  $q_h$  markets will be sustained so long as men and women can coordinate and prostitutes can cover their costs.

---

<sup>17</sup>For most of the remaining discussion, we focus discussion on the BMI and the BMI classification coefficients. The omitted variable when using BMI classifications as controls is respondents with BMI values between 18.5 and 25 ("normal").

### 5.3.1 Craigslist entry, broadband and marginal entrants

Our hypothesis predicts that reductions in  $\psi_t$  will cause new prostitutes to select upon individuals with higher values of  $A$ .<sup>18</sup> We use three variables that we believe plausibly altered the value of  $\psi$  over this period: the dispersion of broadband technology by city and year and the entry of Craigslist into each city over time. The effect of broadband was found to be associated with expanding indoor markets and at some extensive margins, declining street markets (Cunningham and Kendall, 2011*b*). Broadband allowed for improved search time online to coordinate between buyers and sellers through communication channels that did not require public visibility, like street corners. The entry of Craigslist into markets similarly allowed for prostitutes to advertise regularly and safely away from higher risk locations, such as street corners, where the chance of detection was much higher (Cunningham and Kendall, 2011*c*). We report these dates in tables 4 and 5.

Reviewers listed 5 categories,  $k$ , for prostitute’s body types: thin bodies, average bodies, athletic bodies, baby-fat bodies, and fat bodies. We test our theory that declining values of  $\psi$ , proxied by Craigslist entry and increasing proliferation of broadband technology, caused a shift in the relative distribution of body type away from “unattractive body types” and towards more attractive ones using multinomial logit:

$$\text{Prob}(Y_i = j|\mathbf{x}_i) = \frac{e^{\beta_j x_i}}{1 + \sum_{k=1}^5 e^{\beta_k x_k}} \text{ for } j = 1, 2, 4, 5 \quad (20)$$

where  $j$  is the body type and “ $j = 3$ ” is the base category is prostitutes with athletic body types which is our proxy for more ideal body types. We estimate the model using maximum likelihood with robust standard errors. Coefficients were transformed using exponentiation, and therefore each cell is a relative risk ratio. Values less than 1 are negative relative risks equalling 1 minus the coefficient. Values more than 1 are positive

---

<sup>18</sup>First we note some of the weaknesses of TER for these purposes. TER is a non-random sample of prostitutes by city and calendar date (month/year) created by clients of prostitutes self-selecting to review. Changing technological shocks may mean either changes in demand (ie, more attractive males enter the market as probabilities of detection fall), changes in supply (ie, more attractive females enter the market as probabilities of detection fall), or both. While our identification strategy may enable us to see how that particular intermediary’s composition was altered given changes in the probabilities of detection, we cannot say whether the mechanism has been identified given that the technological shocks affects both males and females.

effects.

Conditional on the number of new providers that appeared in a month, as well as women’s characteristics, state characteristics, state and year fixed effects, Craigslist’s opening of its “erotic services” classified section into a city was associated with a 13% reduction in the relative likelihood a prostitute reviewed was classified as “thin” relative to being classified as “athletic”. Likewise, a one standard deviation in broadband penetration was associated with a 26.8% reduction in the probability that a new prostitute reviewed was “thin”.<sup>19</sup> While Craigslist entry was associated with a negative probability a prostitute had an average body type in her review, it was not statistically significant. Again a 1-standard deviation increase broadband was associated with a 47.7% reduction a prostitute reviewed had “average” body type listed with her review. Finally, both Craigslist and its erotic services was associated with declining probability of having “fat” listed in a review, with more precise effects estimated for Craigslist the site (-18.5%) than erotic services (-16.1%).

### 5.3.2 More attractive, higher quality, smaller marketshare

Our hypothesis predicts that more attractive prostitutes will sort into smaller “higher quality” ( $q_h$ ) sectors where prices are higher and by definition clients are willing to travel smaller distances. To examine this, we estimated first the differences in total output by week using the following count model estimated with negative binomial:

$$Y = \alpha + \beta_1 BMI_i + \beta_2 \mathbf{x}_i + \epsilon_i \tag{21}$$

where  $Y$  is the number of non-negative client events she produced in the last week, and  $\epsilon_i$  is distributed poisson with over-dispersion. We correct for over-dispersion with robust standard error correction. Our results are presented in table 10. As can be seen, the higher earnings observed for underweight prostitutes versus obese prostitutes is caused by differences in the number of clients each reported. Underweight prostitutes saw 0.76 more clients in a week, whereas obese prostitutes saw -0.47 fewer clients, than normal

---

<sup>19</sup>Broadband has a standard deviation of 14.9, therefore  $14.9 \times (1 - 0.982) = 0.2682$ .

prostitutes which is consistent with the prediction that normal BMI prostitutes operate within  $x_h$  distances where they see fewer clients as a result and operate as monopolists - both of which reduce the quantity demanded. The higher volume among underweight prostitutes is noteworthy, as it is driven primarily by 1.08 more regulars per week, as well as 0.55 more new clients. Obese prostitutes and normal BMI prostitutes both see fewer clients, particularly fewer regulars, which we will explain in more detail below.

To examine the effect of BMI on prices, we created an average payment based on respondents' last five clients. Mean values for payments were \$611, \$623, \$371 and \$240 for underweight, normal, overweight and obese prostitutes, respectively. We estimated the following model using OLS:

$$Payment = \alpha + \beta_1 BMI + \beta_2 \mathbf{x} + \epsilon \quad (22)$$

where *payment* is the average payment from the previous 5 sessions. These results are shown in table 11. We report our results for both levels and the natural log of payments. We estimate a semi-elasticity of -0.05, essentially identical to what we estimated for weekly earnings in table 9. Interestingly, though underweight BMI is associated with 15 percent lower mean payments compared to women with normal BMI, the difference is not statistically significant. Women with overweight and obese BMI reported average payments that were \$240 – \$430 less than women with normal BMI. We examined other specifications of the payment variable, such as using only the most recent payment in place of the mean payment, and the results were qualitatively the same. Underweight prostitutes payments do not differ statistically from normal BMI in their gross levels or log transformations at the mean, though they are always lower qualitatively.

We examined whether we could find evidence that women who were more attractive were providing services to men with higher willingness to pay. We do this in two indirect ways. First, we examine the length of time spent with clients. Time spent with clients may proxy for “higher quality” if longer time is spent engaging in conversation, foreplay, and quasi-emotional services. Time may also be a dimension along which monopolists bundle sexual services with other services to price discriminate. We estimate the following



models:

$$\ln(\text{length}) = \alpha + \beta_1 \text{BMI} + \beta_2 \mathbf{x} + \epsilon \quad (23)$$

$$\text{Prob}(Y = 1|\mathbf{X}) = \alpha + \beta_1 \text{BMI} + \beta_2 \mathbf{x} + \epsilon \quad (24)$$

. Equation 23 is estimated using OLS with robust standard errors and equation 24 is estimated using probit. We report the marginal effects for each in table 12. On average, higher body mass index values was associated with 2.3 percent reductions in time for every 1-point difference in BMI. Unlike payments, too, we find that underweight, overweight, and obese prostitutes were all more meeting with clients for shorter periods of time with values ranging from 23 to 37 percent. We estimated equation ?? using probit and found that underweight (overweight) prostitutes were 36 percent (29 percent) more likely to spend 60 minutes or less with a client.

Our second test of the hypothesis that women with normal BMI sorted into smaller “high quality” product markets frequented by johns with higher willingness to pay was to estimate the effect of client characteristics on the prices they paid using only the within-prostitute variation in prices and characteristics from the client file. We estimated the following model:

$$\ln(\text{payment}_{t,j}) = \alpha_{t,j} + \beta_1 \mathbf{x}_{t,j} + \rho_t + \epsilon_{t,j} \quad (25)$$

where  $j$  is a particular john, or transaction, and  $t$  is a particular provider. We estimated equation 25 using fixed effects with robust standard errors. The results are presented in table 15. In the first column, we present a combined model in which all body types were included. Because we estimated the model using provider fixed effects, all time-invariant provider characteristics are eliminated. Therefore we estimated both a combined model and models separately using only the BMI category listed in the column header. We focus our results on the normal and overweight results for the moment.

We do not discuss all of the results in this table at this moment since we will note it later when we look for evidence in compensating differentials. For now, we only note the increasing price charged to older males for women with normal BMI versus the declining price charged to older males for overweight prostitutes. Women with normal BMI charged

males, conditional on all observable characteristics of the male and the transaction itself, 2.3 percent more for every 1 year increase in his age. We calculate the maximum age of this age-payment profile to be 55 years of age for women with normal BMI. As we control for provider fixed effects, these cannot be attributed to unobserved provider heterogeneity. Rather, for this one BMI category, there is a systematic tendency to charge higher prices to older males for what is essentially the identical service. We tentatively suggest that this in combination with other findings may be evidence that insofar as older males are men with more income, then they may be men who also have better marital outside options,  $A_j$ . If so, then it is interesting that prostitutes with normal BMI matching with these older males charge them higher prices for what appears to be the identical service (i.e., controlling for all observable client and transaction characteristics, and removing time-invariant provider fixed effects). This may suggest that older males having a higher willingness to pay, because they are unwilling to travel further distances, effectively create market power for more attractive prostitutes. Note that we do not see this same age-payment parabola for other weight groups - in fact, as we will note in hypothesis 3 later, underweight prostitutes in fact price older males *exactly opposite* the pattern observed for women with normal BMI.

## 5.4 Hypothesis 2: Attractiveness and security

Our second hypothesis is drawn from equation 4. Note that the difference in marital payout for more attractive body types is larger than the marital payout for less attractive body types. Consequently, agents with higher  $A$  proxied by normal BMI will be more willing to pay to avoid losing marital options. They will specifically pay up to the difference in marital payout net of the surplus from prostitution. This hypothesis suggests that we should observe prostitutes of normal BMI levels selecting more secure methods, such as avoiding public solicitation methods, exposing themselves to less risk, and pricing risk when possible. We use a variety of evidence to corroborate these predictions.

First we examined the differences in historical risk and victimization BMI category.

We estimate the following probability model using probit:

$$\text{Prob}(Y_k = 1) = \Phi(\alpha + \beta_1 BMI + \beta_2 \mathbf{x} + \epsilon) \text{ for } k = 1, 2 \quad (26)$$

where  $Y$  is dichotomous referring to  $k = 1$  whether a respondent in SASP self-reportedly had ever been assaulted by a client in the past or  $k = 2$  whether she had ever solicited a client from a street location. Both models were estimated separately and reported in table 13. We find that for overweight women, there’s a marginally lower probability of having ever been assaulted ( $p < 0.1$ ), which is not consistent with the thesis presented. But for obese prostitutes, there is 24 percent higher probability of having ever been assaulted by a client ( $p < 0.05$ ). Obese prostitutes were 18.6 percent more likely to have ever attempted to solicit a client from a street location – a form of advertising with a considerably higher risk of detection and arrest given its visibility (Cunningham and Kendall, 2011*b*; Weitzer, 2005).

Obese prostitutes being more likely to advertise ever from streets is consistent with numerous possibilities not contained in our thesis, though. For instance, hypothesis 1 states that there exists a demand curve for prostitution consisting of males willing to travel some maximum distance,  $x_m$ , from their ideal location,  $t^*$ . If female prostitutes service those men, then the residual demand may leave a market demand of men so thin that prostitutes furthest from  $t^*$  must broadcast their identity in order to find a willing client. Higher rates of street solicitation among obese prostitutes may, in other words, simply represent the residual demand curve that obese prostitutes operate on, as opposed to a willingness to use security by johns and prostitutes with higher  $A$  values. A correlation between street prostitution and obesity may also simply represent the endogeneity of obesity to poverty. Therefore, to examine hypothesis 2, we explore other dimensions of risk and managing  $\psi_t$  by BMI.

Prostitutes in SASP were asked first whether they ever used “any” kind of security methods before meeting with a new client, and if so, what kinds of methods did they use. The focus on new clients is important because new clients pose unknown risks. Prostitution is a dangerous crime, particularly for females who are at considerable risk

of violence given the inability to enforce informal contracts with clients and given the asymmetries in strength inherent in the male/female encounter. Therefore, we might expect the willingness to pay for more security to show up in the encounter with a new client.

Table 14 lists along the column headings six different security methods respondents were asked about, which included radio button options as well as open-ended answers that we used to code responses not listed. As prostitutes could use more than 1 of these methods or none of these methods, they weren't mutually exclusive and therefore we did not estimate a multinomial logit. For simplicity, we present simple probit results with transformed coefficients expressed as marginal effects. Outcomes tested were "references", "googles new clients", "required identification", "calls the client at work", "uses gut instincts" and "does a background check". Only two of the outcomes (google and background check) yielded any significant correlation with BMI, and therefore we focus only on these for the sake of brevity. Obese prostitutes were 19.3 percent more likely to use Google to learn about a new client compared to normal BMI respondents, but underweight, overweight and obese prostitutes were all significantly less likely to do a background check than normal BMI which is consistent with the hypothesis from our model (16.2, 7.9 and 9.4 percent less likely, respectively).

A higher use of google to learn about a new client would appear to be consistent with efforts to screen new clients, but in fact given the plethora of other more effective screening options prostitutes can use - such as "references" or background checks - it is worth noting that it's probable that google does not reveal much about the relative risk of a new client. If a new client is using an alias, for instance, then google will not reveal any information. Background checks, on the other hand, require time and money, and will result sometimes in rejecting a new client. Women with normal BMI are systematically more likely to initiate a somewhat costlier form of screening relative to other weight classes before meeting with a new client.

We present one last piece of evidence consistent with our prediction that more attractive agents will pay more to reduce the risk of detection. We return to our fixed effects model, equation 25, in which we estimated the effect of different client and trans-

action characteristics on the payment received. Table 15 presents these results and we focus primarily on the differential prices paid by clients who were regulars versus those who were new clients. Women with normal BMI charged their regular clients 4.7 percent less per payment price than someone purchasing the same service who was a new client, conditional on all observable client and transaction characteristics. None of the other weight classes show statistically significant results for this variable though underweight and overweight prostitutes had qualitatively similar results.

We suggest that this is consistent with the notion that more attractive prostitutes and more attractive clients are willing to pay more to reduce the risk of detection. For female prostitutes, this translated into a risk premium they charge new clients, as well as a tendency to see fewer new clients at all (table 10) and a tendency to screen new clients more thoroughly (table 14). Once the prostitute learns information about her client, she may update her beliefs about his type, and charge him less.

This finding has alternative explanations, though. For instance, one explanation is that prostitutes with normal BMI operate in the quasi-romantic sex markets,  $q_h$ , where repeat business is associated with a feigned attachment to the client. Insofar as these regular clients and the more attractive prostitutes form quasi-romantic client/vendor relationships, then prostitutes may accept lower prices because the cost of the service is considerably lower. We are unable to rule out all the competing hypotheses with these data.

### **5.5 Hypothesis 3: Compensating differentials in the competitive region**

When prostitutes differ with regards to their distances from males' idealized body type and appearance,  $t^*$ , but interact with one another in overlapping monopoly markets because the number of prostitutes in competition is high, then a john located at a point where the two markets overlap will be indifferent between buying from either prostitute if the net surplus from either prostitute is identical (equation 15). When there is only one service the two prostitutes provide, this implies that the less preferred prostitute will charge lower

prices and see fewer clients overall. But if there are two types of services offered, then the john will purchase from the less preferred vendor if the net surplus from the second service exceeds the net surplus of the more preferred service, either because the more preferred service provider has too high of a price for that second service or she refuses to sell the service at all.

We explore this intuition empirically by first noting the results from our earlier negative binomial model presented in table 21. Obese prostitutes saw 0.45 fewer regular clients in a week than normal prostitutes, but underweight prostitutes saw 1.08 more regulars than normal BMI prostitutes. We tested the significance that the underweight and obese results differed from one another, as well. For the first model regarding the total number of clients, our  $\chi^2$  test statistic on the joint significance of our test that the obese and underweight result equalling zero was 12.35 for the overweight and underweight result with a  $p$ -value of 0.002. For the second model regarding the total number of regular clients, the  $\chi^2$  test statistic was 16.98 with  $p$ -value of 0.0002. And for the third model regarding the total number of new clients seen in the previous week, the  $\chi^2$  test statistic was 6.04 with a  $p$ -value of 0.0488. For all types of clients seen - both the total number and the kind - we find that the less attractive prostitutes proxied by obesity saw fewer clients than the underweight prostitutes as well as the normal prostitutes. Obese prostitutes see fewer clients than their underweight rivals in the competitive markets.

Next we examined the differences in prices as well as the provision of different services using fixed effects estimation. We estimate the following model:

$$\ln(\text{payment}) = \beta_1(BMI \times \text{sex act}) + \beta_2\text{sex act} + \beta_3\mathbf{x} + \epsilon \quad (27)$$

Note that because the BMI variable is time-invariant, it is eliminated in the fixed effects strategy. It's interaction, though, is not. We estimated this model using only the within-prostitute data contained in the client file for the combined sample. We estimated eight models so that the readability of the results would be more straightforward for the reader. Table 16 presents these results. While we controlled for all other covariates, since they are similar to what we reported in table 15, we do not reproduce them here.

Note that obese prostitutes are the only group between overweight and obese for whom the interaction is statistically significant, so we focus only on those individuals here. Compared to the rest of the sample, obese prostitutes charged 22.1 percent less for anal sex with a condom. This is consistent with equation 15 in which the marginal john who is indifferent between two prostitutes, one of whom is located further from  $t^*$ , will only purchase from the less preferred prostitute if the price is lower. We find this only for anal sex with a condom.

But perhaps more importantly is the interaction of obesity with anal sex without a condom. Unprotected anal receptive intercourse carries the highest HIV transmission rate of all types of sex acts (Laumann, Gagnon, Michael and Michaels, 1994). Not surprisingly, therefore, we find that the obese prostitutes in the sample when they do choose to provide it charge johns a 58.7 percent higher cost. Note, this estimate is conditional on all observed client and transaction characteristics and nets out all unobserved time-invariant heterogeneity associated with a particular prostitute. Given the higher markup on this service, it may be that obese prostitutes are able to provide it in equilibrium because women closer to  $t^*$  either refuse to provide it given its higher risk or clients are unwilling to pay the corresponding compensating differential associated with it for more beautiful prostitutes.

Finally, we return to a previous table (table 15). As we noted previously, normal prostitutes charged higher prices to older clients, with the peak age-price at 55. But note that overweight prostitutes show a pattern of pricing that is almost exactly opposite. For every 1 year increase in a client's age, overweight prostitutes charged 2.2 percent *lower* prices, with the minimum age at 45. This is a challenging result to completely decipher. We suggest that it is generally consistent with equation 15, though. Younger males may differ from older males in ways that correspond to risk or preferences, and we are unable to separate the two. But, in fielding our survey, several prostitutes noted that typically young and old male clients differed with regards to income as well as preferences. It was more common, we were told, for a young client to come to the meeting intoxicated, and/or to be verbally abusive to the sex worker. They were younger, and therefore oftentimes were rowdier clientele with lower income. Some prostitutes with whom we spoke by telephone

during the fielding of the survey told us that they had a policy *not* to see clients under a particular age threshold. Hence it may be that overweight prostitutes when they see young clients, because these men are priced out of the more attractive markets because of these types of bans, possess some limited market power. This would be an interpretation that is consistent with hypothesis 2.

As male clients age, the equilibrium price falls because older johns can purchase from multiple vendors. Overweight prostitutes would therefore face a male who could purchase from a prostitute with a more ideal body type. He will, according to equation 15, only purchase from the lower quality female if her price is lower. This interpretation is consistent with hypothesis 3.

## 6 Conclusion

Our study has attempted to draw together numerous theories of prostitution and beauty into a single monopolistic competition model. Prostitutes compete in various markets and the returns to beauty in the marriage market ultimately impact the market structure of the commercial sex market via prostitute and john concerns over detection. As the internet has expanded across the United States, underground markets have changed such that the composition of sex workers and clients are increasingly selecting upon more attractive males and females. This shift has in turn potentially altered the overall structure of the market from historical sex markets as new entrants manage risks through greater screening, as well as provide services meant to capture johns with higher preferences for beauty type and appearance.

Our model helps explain and highlight those types of workers most at risk for victimization and STI transmission, as well. We find that obese prostitutes, in particular, face the strongest competition from rivals as a result of the new technologies, and in order to compete, they may engage in high risk activities. Specifically, obese prostitutes lower prices on the anal sex with a condom by 22.1 percent relative to other prostitutes, as well as offer the highest risk acts - anal sex without a condom - at a compensating increase in price of 58.7 percent. This result is in some respects similar to what others have found,



such as Gertler, Shah and Bertozzi (2005) as well as Averett, Corman and Reichman (2012). Gertler, Shah and Bertozzi (2005), as noted earlier, estimated that compensating differentials in developing countries increased as a function of the increasing beauty of the sex worker, but here we find compensating differentials for the *least attractive* sex worker. We believe that this is because more attractive prostitutes in the United States, due to improvements in information technology, sort into markets where such risk-taking is less common potentially because clients could not afford the higher price. Where it occurs in our sample, it is typically among the obese prostitutes who may be doing so in an effort to gain some additional revenue.

**Table 1** Description of provider invariant variables, SASP, 2008-2009

Variables	Mean	Std. Dev	Min	Max	N
Body mass index	23.33	5.80	16	55	532
Underweight (BMI<18.5)	0.10	0.30	0	1	532
Normal (18.5BMI < 25)	0.62	0.49	0	1	532
Overweight (25≤ BMI <30)	0.13	0.34	0	1	532
Obese (BMI≥30)	0.11	0.31	0	1	532
Week's earnings from prostitution	2536.01	3305.81	5	30000	381
Week's earnings including non-workers	1845.55	3037.10	0	30000	524
Ln(earnings)	7.27	1.12	2	10	381
Ln(earnings)	7.27	1.12	2	10	381
Saw clients last week	0.74	0.44	0	1	530
Total clients	5.69	11.40	0	76	527
Regular clients	2.98	6.72	0	45	525
New clients	2.69	5.20	0	31	525
Age	28.12	6.63	18	65	532
Age-squared	834.83	429.96	324	4225	532
Years since began prostitution	5.36	4.90	0	34	532
White	0.63	0.48	0	1	532
Black	0.11	0.32	0	1	532
Other race	0.08	0.27	0	1	532
Less than high school	0.07	0.26	0	1	532
High school	0.10	0.31	0	1	532
Some college	0.43	0.50	0	1	532
College graduate	0.39	0.49	0	1	532
Single, never married	0.46	0.50	0	1	532
Married	0.13	0.34	0	1	532
Widow	0.00	0.07	0	1	532
Cohabiting	0.19	0.39	0	1	532
Divorced	0.17	0.38	0	1	532
Married, separated	0.04	0.21	0	1	532
Parent of a child	0.36	0.48	0	1	532
Less than high school (mother)	0.15	0.36	0	1	532
High school (mother)	0.22	0.41	0	1	532
Some college (mother)	0.22	0.41	0	1	532
College graduate (mother)	0.29	0.46	0	1	532
Post-graduate (mother)	0.10	0.31	0	1	532
Less than high school (father)	0.12	0.33	0	1	532
High school (father)	0.23	0.42	0	1	532
Some college (father)	0.18	0.39	0	1	532
College (father)	0.25	0.43	0	1	532
Post-graduate (father)	0.17	0.37	0	1	532
Inverse Mills Ratio	2.71	0.29	2	3	532
Concern of police detection	4.01	2.38	1	10	520
Concern of family detection	4.00	3.06	1	10	521
Chance of HIV infection	2.68	2.05	1	10	521
Concern of violent client	4.07	2.59	1	10	522
Ever assaulted by client	0.14	0.35	0	1	524
Solicited from a street before	0.15	0.35	0	1	530
Takes precautions with new clients	0.79	0.41	0	1	516
Requires third-party references with new clients	0.60	0.49	0	1	532
Uses search engines to screen new clients	0.57	0.50	0	1	532
Requires state identification with new clients	0.32	0.47	0	1	532
Conducts background check on new clients	0.21	0.41	0	1	532
Average session length	128.72	268.29	5	3792	515
Ln(average length)	4.39	0.76	2	8	515
Session shorter than hour	0.28	0.45	0	1	515
Session 60 min or longer	0.72	0.45	0	1	515
Recent client was white	0.83	0.38	0	1	532
Recent client was black	0.02	0.15	0	1	532
Recent client other race	0.09	0.28	0	1	532

Data was collected from August 2008 to June 2009. Census division dummies and month dummies were also collected but are not shown. Inverse probability weights are used to report summary statistics.

**Table 2** Description of provider variant session characteristics, SASP, 2008-2009

Variables	Mean	Std. Dev	Min	Max	N
Hourly adjusted price	346.21	329.41	50	7000	1969
Ln(Wage)	5.65	0.57	4	9	1969
Total cost of session	496.13	1115.63	25	28000	1969
Ln(Payment)					
Session length (min.)	119.75	366.54	5	7200	1969
Session < 60 min	0.28	0.45	0	1	1969
Session ≥ 60 min	0.72	0.45	0	1	1969
Age of client	43.03	10.88	18	91	1969
Age of client squared	1969.61	1007.64	324	8281	1969
Regular client	0.54	0.50	0	1	1969
Asian client	0.07	0.26	0	1	1969
Black client	0.05	0.23	0	1	1969
White client	0.81	0.39	0	1	1969
Other race client	0.14	0.34	0	1	1969
Second provider present at session	0.06	0.23	0	1	1969
Kissed client	0.59	0.49	0	1	1969
French kissed client	0.55	0.50	0	1	1969
Met client at hotel room	0.47	0.50	0	1	1969
Traveled out of state for assignation	0.16	0.37	0	1	1969
Provided massage to client	0.36	0.48	0	1	1969
Cunnilingus	0.58	0.49	0	1	1969
Vaginal sex not provided	0.25	0.44	0	1	1969
Vaginal sex with condom	0.68	0.47	0	1	1969
Vaginal sex, no condom	0.06	0.23	0	1	1969
Fellatio not provided	0.18	0.39	0	1	1969
Fellatio with condom	0.29	0.46	0	1	1969
Fellatio, no condom	0.51	0.50	0	1	1969
Anal sex not provided	0.94	0.24	0	1	1969
Anal sex, condom	0.05	0.21	0	1	1969
Anal sex, no condom	0.01	0.11	0	1	1969
First introduction to client was word-of-mouth	0.01	0.09	0	1	1969
First introduction to client was over the phone	0.31	0.46	0	1	1969
First introduction to client was through referral	0.04	0.20	0	1	1969
First introduction to client was through public face-to-face	0.02	0.12	0	1	1969
First introduction to client was by email/Internet	0.58	0.49	0	1	1969
First introduction occurred other way	0.04	0.20	0	1	1969
Ln(payment)	5.76	0.76	3	10	1969

Data was collected from August 2008 to June 2009. Census division dummies and month dummies were also collected but are not shown. Inverse probability weights are used to report summary statistics.

**Table 3** Description of female respondent characteristics, GSS, 1991-2010

Variables	Mean	Std. Dev	Min	Max	N
Ever received compensation for sex	0.02	0.13	0	1	12886
Age	45.25	17.02	18	89	12886
Age-squared	2337.21	1717.89	324	7921	12886
Less than high school	0.14	0.34	0	1	12886
High school	0.54	0.50	0	1	12886
Some college	0.08	0.27	0	1	12886
College graduate	0.24	0.43	0	1	12886
White	0.79	0.40	0	1	12886
Black	0.14	0.35	0	1	12886
Other race	0.06	0.25	0	1	12886
Married	0.49	0.50	0	1	12886
Widow	0.11	0.31	0	1	12886
Divorced	0.15	0.36	0	1	12886
Married, separated	0.04	0.19	0	1	12886
Single, never married	0.21	0.41	0	1	12886
Parent of a child	0.76	0.43	0	1	12886
Less than high school (mother)	0.31	0.46	0	1	12886
High school (mother)	0.44	0.50	0	1	12886
Some college (mother)	0.04	0.19	0	1	12886
College graduate (mother)	0.08	0.27	0	1	12886
Post-graduate (mother)	0.03	0.17	0	1	12886
Less than high school (father)	0.30	0.46	0	1	12886
High school (father)	0.31	0.46	0	1	12886
Some college/vocational (father)	0.02	0.14	0	1	12886
Bachelors (father)	0.08	0.27	0	1	12886
Post-graduate studies (father)	0.05	0.22	0	1	12886

General Social Survey (GSS) waves 1991, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010 were used in estimation. Census division variables were used in estimation of the selection equation but are not shown. Summary statistics are weighted by household size sampling weights.

**Table 4** Dates When City-Specific Craigslist Boards Launched

Date of Introduction		City	TER Escorts (all years)	Cumulative Share of TER Escorts
Aug	1998	San Francisco	5,945	7.69%
Aug	2000	Boston	2,822	11.34%
Feb	2001	Chicago	3,960	16.47%
		Los Angeles	9,151	28.30 %
		New York	7,682	38.24%
		Portland	738	39.20%
		San Diego	2,093	41.91%
		Seattle	1,582	43.95%
		Washington, DC	4,980	50.40%
Apr	2001	Atlanta	3,634	55.10%
		Denver	1,455	56.98%
		Vancouver	842	58.07%
Jun	2001	Austin	425	58.62%
Nov	2002	Miami	2,809	62.25%
		Minnesota	1,857	64.65%
		Philadelphia	1,428	66.50%
		Phoenix	2,313	69.49%
Apr	2003	Dallas	2,073	72.18%
		Detroit	1,141	73.65%
		London	2,238	76.55%
		Toronto	1,903	79.01%
May	2003	Houston	1,667	81.17%
Nov	2003	Carolinas	1,704	83.37%
		Cleveland	821	84.43%
		Hawaiian Islands	296	84.82%
		New Orleans	280	85.18 %
		Orlando	312	85.58%
		Tampa	1,094	87.00%
Dec	2003	Montreal	806	88.04%
Feb	2004	Las Vegas	1,962	90.58%
		Nashville	194	90.83%
Mar	2004	Columbus	415	91.37%
Apr	2004	Indiana	284	91.73%
Sep	2004	Gold Coast, CA	184	91.97%
		Jacksonville	11	91.99%
		New Mexico	62	92.07%
		Salt Lake City	391	92.57%
		Tucson	113	92.72%
Nov	2004	Orange County	1,762	95.00%
		Reno	281	95.36%
Mar	2005	New Jersey	2,611	98.74%
Jun	2005	New England	687	99.63%
		Tijuana	214	99.90%
Feb	2006	Palm Springs	74	100.00%
Total			77,295	100.00%

**Table 5** Dates When City-Specific Craigslist “Erotic Services” Boards Launched

Date of Introduction		City	TER Escorts (all years)	Cumulative Share of TER Escorts
Nov	2002	San Francisco	5,945	7.69%
Jul	2003	Boston	2,822	11.34%
		Los Angeles	9,151	23.18%
		New York	7,582	33.12%
Sep	2003	Phoenix	2,313	36.11%
		San Diego	2,093	38.82%
		Seattle	1,582	40.87%
		Washington, DC	4,980	47.31%
Oct	2003	Austin	425	47.86%
		Chicago	3,960	52.98%
		Dallas	2,073	55.66%
		Denver	1,455	57.55%
		Detroit	1,141	59.02%
		Houston	1,667	61.18%
		London	2,238	64.08%
		Miami	2,809	67.71%
		Minnesota	1,857	70.11%
		Philadelphia	1,428	71.96%
		Portland	738	72.91%
		Toronto	1,903	75.38%
		Vancouver	842	76.47%
Nov	2003	Atlanta	3,634	81.17%
		Carolinas	1,704	83.37%
		Cleveland	821	84.43 %
		Hawaiian Islands	296	84.82%
		New Orleans	280	85.18%
		Orlando	312	85.58%
		Tampa	1,094	87.00%
Dec	2003	Montreal	806	88.04%
Feb	2004	Las Vegas	1,962	90.58%
		Nashville	194	90.83%
Mar	2004	Columbus	415	91.37%
Apr	2004	Indiana	284	91.73%
Sep	2004	Gold Coast, CA	183	91.97%
		Jacksonville	11	91.99%
		New Mexico	62	92.07%
		Salt Lake City	391	92.57%
		Tucson	113	92.72%
Nov	2004	Orange County	1,762	95.00 %
		Reno	281	95.36%
Mar	2005	New Jersey	2,611	98.74%
Jun	2005	New England	687	99.63%
		Tijuana	214	99.90%
Feb	2006	Palm Springs	74	100.00%
Total			77,295	100.00%

**Table 6** Relative risk ratios of broadband dispersion and Craigslist entry on the probability a new prostitute reviewed had a particular body type, TER, 1999-2007

	Prob=Thin	Prob=Average	Prob=Athletic	Prob=Babyfat	Prob=Fat
Craigslist erotic services entry	0.870** (0.054)	0.975 (0.068)	1.000 (.)	0.995 (0.078)	0.839* (0.084)
Craigslist entry	1.016 (0.061)	0.954 (0.064)	1.000 (.)	0.897 (0.068)	0.815** (0.078)
Broadband lines per household	0.982** (0.008)	0.968*** (0.009)	1.000 (.)	0.989 (0.010)	0.985 (0.013)
Total new providers	1.000 (0.001)	1.000 (0.001)	1.000 (.)	0.999 (0.001)	1.000 (0.001)
White	1.144*** (0.056)	1.287*** (0.072)	1.000 (.)	1.215*** (0.078)	1.305*** (0.105)
Black	0.715*** (0.043)	0.715*** (0.051)	1.000 (.)	0.905 (0.070)	1.281** (0.124)
Asian	2.403*** (0.136)	1.903*** (0.124)	1.000 (.)	1.525*** (0.118)	0.986 (0.106)
Hispanic	0.873** (0.049)	1.334*** (0.083)	1.000 (.)	1.575*** (0.110)	1.465*** (0.133)
18-20 year old	1.778*** (0.099)	0.742*** (0.047)	1.000 (.)	1.228*** (0.085)	0.342*** (0.035)
21-25 year old	1.279*** (0.055)	0.660*** (0.029)	1.000 (.)	0.877** (0.045)	0.367*** (0.023)
26-30 year old	1.011 (0.046)	0.748*** (0.035)	1.000 (.)	0.838*** (0.045)	0.504*** (0.031)
36-40 year old	1.313*** (0.097)	1.492*** (0.107)	1.000 (.)	1.080 (0.095)	1.775*** (0.149)
41-45 year old	1.512*** (0.170)	1.619*** (0.175)	1.000 (.)	1.463*** (0.186)	3.149*** (0.358)
46 and older	1.834*** (0.305)	2.127*** (0.341)	1.000 (.)	1.852*** (0.343)	5.135*** (0.825)
A cup	10.119*** (0.603)	1.280*** (0.102)	1.000 (.)	0.528*** (0.052)	0.387*** (0.052)
B cup	3.926*** (0.139)	1.440*** (0.055)	1.000 (.)	0.532*** (0.024)	0.344*** (0.021)
C cup	1.714*** (0.059)	1.236*** (0.043)	1.000 (.)	0.666*** (0.024)	0.398*** (0.019)
Meth purity	0.998 (0.004)	0.996 (0.004)	1.000 (.)	1.003 (0.005)	0.983*** (0.006)
Meth price	1.000* (0.000)	1.000* (0.000)	1.000 (.)	1.000 (0.000)	1.000 (0.000)
Cocaine purity	0.988 (0.030)	1.021 (0.035)	1.000 (.)	1.004 (0.039)	1.042 (0.051)
Cocaine price	0.999 (0.001)	1.000 (0.001)	1.000 (.)	1.000 (0.002)	1.000 (0.002)
Heroin purity	0.990 (0.007)	0.995 (0.008)	1.000 (.)	0.999 (0.009)	1.002 (0.011)
Heroin price	1.000 (0.000)	1.000 (0.000)	1.000 (.)	1.000 (0.000)	1.001** (0.000)
Policemen per capita	1.000 (0.000)	1.001** (0.000)	1.000 (.)	1.000 (0.000)	1.000 (0.000)
Prisoners per capita	1.000 (0.001)	1.000 (0.001)	1.000 (.)	1.001 (0.001)	1.000 (0.001)
Population density	1.551** (0.324)	0.956 (0.222)	1.000 (.)	0.885 (0.242)	1.476 (0.479)
Real income	1.000 (0.000)	1.000 (0.000)	1.000 (.)	1.000 (0.000)	1.000* (0.000)
State unemployment rate	0.921* (0.044)	0.937 (0.050)	1.000 (.)	1.006 (0.061)	0.909 (0.069)
Pseudo R-squared	0.08				
N	47094				

Each column represents a separate relative probability estimated from a single multinomial logit model. Coefficients were transformed into marginal effects for ease of interpretation. All models include city fixed effects, month-of-year fixed effects and year fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 7** Marginal effects from probit estimation of whether received compensation for sex, female respondents, GSS 1991-2010

Ever received compensation for sex	
Age	0.001*** (0.000)
Age-squared	-0.000*** (0.000)
Black (d)	0.028*** (0.005)
Other race (d)	-0.005 (0.003)
Less than high school (d)	0.010*** (0.004)
Some college (d)	-0.007** (0.003)
College (d)	-0.006** (0.002)
Married (d)	-0.004 (0.003)
Widow (d)	0.003 (0.005)
Divorced (d)	0.003 (0.004)
Separated (d)	0.003 (0.005)
Parent of at least one child (d)	0.001 (0.003)
Less than high school (mother) (d)	-0.006*** (0.002)
Some college/vocational (mother) (d)	-0.009*** (0.003)
Bachelors (mother) (d)	-0.003 (0.003)
Post-graduate studies (mother) (d)	-0.002 (0.005)
Less than high school (father) (d)	-0.003 (0.002)
Some college/vocational (father) (d)	-0.004 (0.006)
Bachelors (father) (d)	-0.002 (0.004)
Post-graduate studies (father) (d)	0.006 (0.006)
1993 (d)	0.005 (0.007)
1994 (d)	0.011 (0.008)
1996 (d)	0.006 (0.007)
1998 (d)	0.008 (0.008)
2000 (d)	0.006 (0.007)
2002 (d)	0.011 (0.008)
2004 (d)	0.012 (0.009)
2006 (d)	0.009 (0.008)
2008 (d)	0.022* (0.012)
2010 (d)	0.010 (0.008)
Pseudo R-squared	0.08
N	12886
Peak of age-prostitute parabola	45.37
Mean of dependent variable	0.02

We estimated the selection equation using probit and calculated the inverse mills ratio using the predicted values. Coefficients have been transformed into marginal effects for ease of interpretation. General Social Survey waves 1991, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010 were used in estimation. The dependent variable in each model is dichotomous variable equalling 1 if female respondent had ever received compensation for sexual intercourse, of which 253 female respondents of the 12,918 answered in the affirmative. Heteroskedastic robust standard errors in parenthesis. All models use household size sampling weights as analytical weights and include Census division fixed effects and year fixed effects. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01



**Table 8** Estimated marginal effects of background characteristics and BMI clinical thresholds, SASP, probit

	Estimate probability of BMI category			
	Underweight	Normal	Overweight	Obese
Age	-0.007 (0.019)	0.002 (0.038)	-0.014 (0.019)	0.026* (0.014)
Age-squared	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000* (0.000)
Years since began prostitution	0.000 (0.003)	-0.003 (0.005)	0.002 (0.002)	-0.001 (0.002)
Black (d)		0.133 (0.272)	-0.116*** (0.041)	0.240 (0.310)
Other Race (d)		0.055 (0.123)	0.024 (0.079)	0.090 (0.082)
Less than high school (d)	-0.054* (0.029)	0.082 (0.188)	-0.022 (0.086)	0.019 (0.090)
Some college (d)	-0.132* (0.077)	-0.019 (0.167)	0.150 (0.105)	0.131 (0.088)
College grad (d)	-0.160** (0.067)	-0.015 (0.144)	0.104 (0.084)	0.158 (0.097)
Married, Cohabiting (d)	-0.075*** (0.028)	0.040 (0.112)	0.118 (0.079)	-0.022 (0.034)
Widow (d)		0.335*** (0.053)	-0.049 (0.054)	
Divorced (d)	-0.036 (0.029)	0.105 (0.091)	-0.035 (0.039)	0.018 (0.041)
Separated (d)	-0.037 (0.034)	0.140 (0.125)	0.041 (0.093)	-0.018 (0.040)
Unmarried, Cohabiting (d)	-0.055** (0.026)	0.149 (0.096)	-0.044 (0.037)	0.021 (0.049)
Parent of at least one child (d)	-0.005 (0.035)	-0.042 (0.076)	0.028 (0.041)	-0.010 (0.030)
Less than high school (mother) (d)	-0.049 (0.055)	-0.073 (0.171)	-0.013 (0.068)	0.105 (0.095)
Some college (mother) (d)	0.090 (0.215)	-0.116 (0.272)	0.102 (0.168)	0.045 (0.124)
College (mother) (d)	0.114 (0.080)	-0.073 (0.106)	0.035 (0.052)	0.016 (0.042)
Post-graduate (mother) (d)	0.055 (0.078)	0.071 (0.110)	-0.019 (0.049)	-0.022 (0.037)
Less than high school (father) (d)	0.531*** (0.201)	-0.444*** (0.122)	0.008 (0.075)	0.116 (0.100)
Some college (father) (d)	-0.015 (0.054)	-0.124 (0.124)	0.175** (0.087)	0.012 (0.048)
College (father) (d)	0.029 (0.058)	-0.107 (0.101)	0.103* (0.059)	-0.023 (0.031)
Post-graduate (father) (d)	0.072 (0.080)	-0.197 (0.148)	0.074 (0.077)	0.076 (0.090)
Inverse Mills Ratio	0.001 (0.272)	0.317 (0.603)	-0.412 (0.297)	0.064 (0.217)
Pseudo R-squared	0.36	0.09	0.13	0.21
N	421	532	532	519
Peak of age-BMI parabola	50.72	249.66	36.96	41.54
Mean of dependent variable	0.14	0.62	0.13	0.11

The dependent variables are dichotomous values of 0 or 1 depending on whether the respondent's BMI is within the clinical range of the column headers (e.g., obesity is BMI  $\geq 30$ ). Models were estimated with probit and marginal effects are reported. Heteroskedastic robust standard errors in parenthesis. SASP inverse probability weights were used in estimation, and all models include Census division fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 9** Estimated effect of bodyweight on natural log of weekly earnings, SASP, OLS

	Ln(Earnings)	
Body Mass Index	-0.050*** (0.011)	
BMI < 18.5		0.843*** (0.321)
25 ≤ BMI < 29.9		-0.370 (0.261)
BMI ≥ 30		-0.697*** (0.165)
Age	0.059 (0.101)	0.034 (0.096)
Age-squared	-0.001 (0.001)	-0.001 (0.001)
Years since began prostitution	-0.000 (0.011)	0.004 (0.011)
Black	0.318 (0.819)	0.323 (0.788)
Other Race	0.231 (0.268)	0.313 (0.279)
Less than high school	-0.150 (0.552)	0.001 (0.527)
Some college	-0.711 (0.487)	-0.529 (0.442)
College grad	-0.305 (0.448)	-0.071 (0.409)
Married, Cohabiting	0.020 (0.278)	0.018 (0.263)
Widow	-0.599* (0.357)	-0.493 (0.360)
Divorced	-0.097 (0.259)	-0.066 (0.253)
Separated	0.200 (0.368)	0.228 (0.375)
Unmarried, Cohabiting	0.023 (0.275)	0.022 (0.270)
Parent of at least one child	0.123 (0.200)	0.177 (0.197)
Less than high school (mother)	0.257 (0.401)	0.242 (0.393)
Some college (mother)	-0.012 (0.592)	0.022 (0.575)
College (mother)	0.017 (0.277)	0.043 (0.261)
Post-graduate (mother)	-0.127 (0.475)	-0.108 (0.454)
Less than high school (father)	0.126 (0.263)	0.092 (0.273)
Some college (father)	-0.127 (0.323)	-0.071 (0.319)
College (father)	0.024 (0.221)	-0.006 (0.219)
Post-graduate (father)	0.031 (0.365)	-0.060 (0.336)
Inverse Mills Ratio	0.423 (1.402)	0.300 (1.385)
R-squared	0.19	0.22
N	381	381
Peak of age-earnings parabola	30.09	25.32
Mean of dependent variable	7.27	7.27

We estimated the effect of bodytype on the natural log of earnings using OLS for BMI and BMI clinical thresholds, respectively. Weekly earnings corresponds only to earnings from prostitution and excludes non-monetary gifts, such as fur coats or diamonds. Heteroskedastic robust standard errors in parenthesis. SASP inverse probability weights are used in estimation, and all models include Census division fixed effects. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 10** Estimated effect of bodyweight on number of clients seen in a week, SASP, negative binomial count models

	Total clients		Regulars		New clients	
Body Mass Index	-0.027** (0.011)		-0.044*** (0.012)		-0.010 (0.013)	
BMI < 18.5		0.771** (0.324)		1.080*** (0.335)		0.545* (0.330)
25 ≤ BMI < 29.9		-0.039 (0.189)		0.052 (0.189)		-0.013 (0.231)
BMI ≥ 30		-0.465** (0.224)		-0.450** (0.209)		-0.363 (0.261)
Age	0.021 (0.092)	0.021 (0.088)	-0.026 (0.092)	-0.035 (0.084)	0.043 (0.103)	0.049 (0.102)
Age-squared	-0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	-0.000 (0.001)	-0.001 (0.001)
Years since began prostitution	-0.011 (0.011)	-0.010 (0.011)	0.004 (0.012)	0.006 (0.012)	-0.022 (0.014)	-0.023 (0.014)
Black	-0.361 (0.801)	-0.344 (0.776)	-1.368* (0.742)	-1.315* (0.699)	0.433 (0.925)	0.452 (0.917)
Other Race	-0.252 (0.321)	-0.153 (0.316)	0.022 (0.299)	0.168 (0.296)	-0.396 (0.384)	-0.311 (0.383)
Less than high school	-0.762 (0.523)	-0.543 (0.513)	-0.656 (0.461)	-0.315 (0.452)	-1.032 (0.629)	-0.885 (0.630)
Some college	-1.119*** (0.429)	-0.880** (0.428)	-0.512 (0.424)	-0.195 (0.422)	-1.744*** (0.474)	-1.552*** (0.486)
College grad	-1.000*** (0.380)	-0.754* (0.387)	-0.318 (0.364)	0.049 (0.377)	-1.630*** (0.428)	-1.449*** (0.446)
Married, Cohabiting	-0.109 (0.288)	-0.070 (0.281)	0.056 (0.299)	0.093 (0.289)	-0.264 (0.328)	-0.212 (0.323)
Widow	-0.873 (0.771)	-0.713 (0.714)	-0.768 (0.784)	-0.443 (0.649)	-1.300 (0.980)	-1.174 (0.918)
Divorced	-0.377 (0.255)	-0.306 (0.245)	-0.418* (0.242)	-0.339 (0.222)	-0.371 (0.295)	-0.298 (0.295)
Separated	0.089 (0.335)	0.118 (0.326)	0.070 (0.328)	0.094 (0.311)	0.239 (0.423)	0.282 (0.429)
Unmarried, Cohabiting	-0.284 (0.260)	-0.268 (0.255)	-0.325 (0.236)	-0.287 (0.216)	-0.185 (0.303)	-0.182 (0.304)
Parent of at least one child	0.337** (0.166)	0.342** (0.165)	0.237 (0.161)	0.255 (0.159)	0.474** (0.208)	0.467** (0.206)
Less than high school (mother)	-0.301 (0.431)	-0.310 (0.408)	0.259 (0.388)	0.215 (0.358)	-0.642 (0.494)	-0.638 (0.485)
Some college (mother)	-0.756 (0.648)	-0.708 (0.615)	-0.133 (0.616)	-0.133 (0.554)	-1.111 (0.722)	-1.060 (0.707)
College (mother)	0.124 (0.278)	0.128 (0.265)	0.209 (0.241)	0.168 (0.229)	0.144 (0.343)	0.148 (0.338)
Post-graduate (mother)	0.208 (0.317)	0.288 (0.305)	0.422 (0.291)	0.515* (0.268)	0.106 (0.385)	0.158 (0.381)
Less than high school (father)	0.296 (0.334)	0.208 (0.334)	0.617* (0.324)	0.468 (0.331)	0.006 (0.361)	-0.049 (0.367)
Some college (father)	-0.603** (0.254)	-0.577** (0.250)	-0.444* (0.239)	-0.436* (0.238)	-0.688** (0.312)	-0.645** (0.312)
College (father)	-0.219 (0.243)	-0.255 (0.231)	-0.041 (0.210)	-0.135 (0.194)	-0.438 (0.292)	-0.428 (0.288)
Post-graduate (father)	-0.447 (0.325)	-0.534* (0.306)	-0.602** (0.296)	-0.763*** (0.264)	-0.344 (0.398)	-0.379 (0.390)
Inverse Mills Ratio	1.399 (1.534)	1.270 (1.496)	-0.101 (1.434)	-0.217 (1.368)	2.580 (1.679)	2.454 (1.666)
Pseudo R-squared						
N	527	527	525	525	525	525
Peak of age-work parabola	55.22	49.09	38.33	42.77	48.12	45.89
Mean of dependent variable	5.69	5.69	2.98	2.98	2.69	2.69

We estimated the correlation between BMI and BMI clinical thresholds and the number of clients seen in a week using the negative binomial count model. We looked at the total clients, the number of regulars and the number of new clients the respondent saw the previous week. We include zero counts for having not worked that week at all. Heteroskedastic robust standard errors in parenthesis. Coefficients have been transformed into marginal effects for ease of interpretation. SASP inverse probability weights are used in estimation, and all models include Census division fixed effects corresponding to the respondent's base of operations. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 11** Marginal effect of bodyweight on payment and ln(payment) received in last session, SASP, OLS

	Payment		ln(payment)	
Body Mass Index	-25.069*** (6.005)		-0.041*** (0.007)	
BMI < 18.5		-115.065 (180.487)		-0.151 (0.148)
25 ≤ BMI < 29.9		-239.892** (98.145)		-0.436*** (0.118)
BMI ≥ 30		-430.360*** (107.222)		-0.673*** (0.107)
Age	-46.841 (46.940)	-53.257 (48.133)	0.026 (0.045)	0.013 (0.042)
Age-squared	0.362 (0.540)	0.429 (0.549)	-0.000 (0.001)	-0.000 (0.001)
Years since began prostitution	-0.535 (7.555)	1.196 (7.673)	-0.004 (0.006)	-0.001 (0.006)
Black	-740.765* (428.911)	-770.787* (451.090)	-0.049 (0.361)	-0.111 (0.355)
Other Race	169.707 (107.875)	158.199 (107.004)	0.148 (0.134)	0.131 (0.133)
Less than high school	-352.224* (188.674)	-335.385* (197.830)	0.061 (0.185)	0.088 (0.192)
Some college	456.314* (235.016)	490.887** (242.974)	0.350* (0.204)	0.412** (0.203)
College grad	559.454** (254.605)	595.228** (259.416)	0.480*** (0.178)	0.542*** (0.181)
Married, Cohabiting	241.102* (126.152)	239.264* (124.795)	0.271* (0.140)	0.275* (0.141)
Widow	-144.991 (287.481)	-141.290 (294.590)	0.114 (0.354)	0.126 (0.377)
Divorced	-16.859 (114.206)	9.415 (113.917)	0.083 (0.101)	0.123 (0.098)
Separated	-317.721** (136.150)	-300.357** (142.040)	-0.215 (0.148)	-0.185 (0.161)
Unmarried, Cohabiting	-95.720 (186.883)	-76.456 (189.050)	-0.067 (0.122)	-0.037 (0.120)
Parent of at least one child	-34.731 (73.052)	-37.981 (74.089)	-0.056 (0.084)	-0.058 (0.085)
Less than high school (mother)	384.484** (170.375)	372.087** (169.922)	0.284* (0.149)	0.264* (0.151)
Some college (mother)	783.080** (343.228)	820.915** (356.692)	0.318 (0.271)	0.383 (0.265)
College (mother)	121.110 (84.196)	160.089* (94.912)	0.095 (0.114)	0.158 (0.115)
Post-graduate (mother)	440.403 (276.258)	467.370* (282.918)	0.399** (0.158)	0.444*** (0.165)
Less than high school (father)	182.444 (122.134)	240.718* (132.791)	0.105 (0.131)	0.194 (0.125)
Some college (father)	217.656 (136.887)	216.657 (139.700)	0.205 (0.137)	0.210 (0.137)
College (father)	154.134 (104.361)	155.983 (105.590)	0.107 (0.111)	0.115 (0.109)
Post-graduate (father)	288.317 (189.212)	282.428 (190.472)	0.329** (0.141)	0.320** (0.138)
Inverse Mills Ratio	-1544.731* (816.592)	-1583.630* (845.777)	-0.374 (0.664)	-0.460 (0.642)
R-squared	0.16	0.16	0.30	0.30
Pseudo R-squared				
N	498	498	498	498
Peak of age-payment parabola	64.64	62.07	29.03	21.59
Mean of dependent variable	538.83	538.83	5.89	5.89

Dependent variable is the most recent payment from a client expressed in levels and logs. OLS coefficients each measure the marginal effects. Heteroskedastic robust standard errors are shown in parenthesis. SASP inverse probability weights are used in estimation, and all models include the provider's primary Census division residential location fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 12** Marginal effect of bodyweight on average session length, SASP, OLS and probit

	Ln(average length)		60 min or longer	
Body Mass Index	-0.023*** (0.006)		-0.009** (0.004)	
BMI < 18.5 (d)		-0.287* (0.170)		-0.363*** (0.117)
25 ≤ BMI < 29.9 (d)		-0.273** (0.109)		-0.287*** (0.084)
BMI ≥ 30 (d)		-0.372*** (0.109)		-0.076 (0.094)
Age	0.032 (0.042)	0.025 (0.040)	0.086** (0.036)	0.079** (0.032)
Age-squared	-0.000 (0.001)	-0.000 (0.001)	-0.001* (0.000)	-0.001* (0.000)
Years since began prostitution	-0.015*** (0.006)	-0.014** (0.006)	-0.010** (0.004)	-0.009** (0.004)
Black (d)	0.547 (0.389)	0.489 (0.381)	0.385*** (0.047)	0.366*** (0.045)
Other Race (d)	-0.199 (0.145)	-0.233 (0.143)	-0.079 (0.145)	-0.122 (0.145)
Less than high school (d)	0.084 (0.223)	0.083 (0.224)	0.272*** (0.032)	0.267*** (0.033)
Some college (d)	0.179 (0.224)	0.203 (0.217)	-0.210 (0.161)	-0.189 (0.150)
College grad (d)	0.328* (0.189)	0.343* (0.189)	0.001 (0.135)	0.002 (0.131)
Married, Cohabiting (d)	0.223 (0.144)	0.226 (0.142)	-0.010 (0.108)	0.006 (0.101)
Widow (d)	-0.236 (0.187)	-0.269 (0.192)	0.052 (0.180)	0.006 (0.201)
Divorced (d)	0.266* (0.144)	0.292** (0.143)	0.165*** (0.054)	0.180*** (0.050)
Separated (d)	-0.005 (0.159)	0.013 (0.165)	0.111 (0.094)	0.125 (0.088)
Unmarried, Cohabiting (d)	0.010 (0.118)	0.020 (0.117)	0.097 (0.067)	0.093 (0.066)
Parent of at least one child (d)	-0.106 (0.088)	-0.111 (0.089)	-0.032 (0.065)	-0.042 (0.066)
Less than high school (mother) (d)	0.266 (0.173)	0.241 (0.165)	-0.172 (0.172)	-0.208 (0.161)
Some college (mother) (d)	-0.149 (0.307)	-0.110 (0.290)	-0.653*** (0.213)	-0.634*** (0.197)
College (mother) (d)	0.047 (0.132)	0.093 (0.127)	-0.070 (0.098)	-0.035 (0.089)
Post-graduate (mother) (d)	0.277 (0.169)	0.307* (0.173)	0.094 (0.086)	0.105 (0.084)
Less than high school (father) (d)	0.155 (0.169)	0.216 (0.160)	-0.061 (0.129)	-0.018 (0.106)
Some college (father) (d)	0.114 (0.141)	0.115 (0.139)	-0.104 (0.114)	-0.082 (0.106)
College (father) (d)	-0.015 (0.102)	-0.004 (0.100)	-0.026 (0.084)	0.010 (0.077)
Post-graduate (father) (d)	0.360** (0.148)	0.362** (0.141)	0.200*** (0.063)	0.206*** (0.062)
Inverse Mills Ratio	0.670 (0.727)	0.615 (0.690)	1.684*** (0.576)	1.597*** (0.524)
R-squared	0.26	0.26		
Pseudo R-squared			0.20	0.23
N	515	515	515	515
Peak of age-length parabola	61.96	66.80	50.25	51.15
Mean of dependent variable	4.39	4.39	0.72	0.72

The first model examines the role of bodytype on the natural log of the length of a particular session (measured in minutes) using OLS. The second model examines the probability the session was 60 minutes or longer using probit. Probit coefficients and OLS coefficients each measure the marginal effects. Heteroskedastic robust standard errors are shown in parenthesis. SASP inverse probability weights are used in estimation, and all models include the provider's primary Census division residential location fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 13** Estimated marginal effect of bodyweight on historical victimization and street solicitation, SASP, probit

	Assaulted ever		Former streetwalker	
Body Mass Index	0.009***		0.007**	
	(0.003)		(0.003)	
BMI < 18.5 (d)		-0.069		0.161
		(0.048)		(0.108)
25 ≤ BMI < 29.9 (d)		-0.054*		-0.001
		(0.030)		(0.040)
BMI ≥ 30 (d)		0.236**		0.186*
		(0.104)		(0.109)
Age	0.031	0.028	-0.046**	-0.041**
	(0.021)	(0.021)	(0.021)	(0.017)
Age-squared	-0.001*	-0.000*	0.000*	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Years since began prostitution	0.008***	0.008***	0.017***	0.016***
	(0.003)	(0.003)	(0.004)	(0.003)
Black (d)	-0.047	-0.075	-0.094**	-0.083**
	(0.113)	(0.079)	(0.038)	(0.038)
Other Race (d)	-0.058	-0.056	0.192	0.209
	(0.039)	(0.038)	(0.138)	(0.148)
Less than high school (d)	0.011	-0.032	-0.008	0.012
	(0.111)	(0.078)	(0.095)	(0.099)
Some college (d)	0.056	0.050	0.047	0.046
	(0.091)	(0.093)	(0.092)	(0.085)
College grad (d)	0.033	0.010	0.020	0.023
	(0.084)	(0.082)	(0.070)	(0.069)
Married, Cohabiting (d)	0.025	0.043	0.070	0.078
	(0.075)	(0.078)	(0.080)	(0.077)
Divorced (d)	0.155*	0.118	0.017	0.013
	(0.087)	(0.075)	(0.064)	(0.059)
Separated (d)	0.143	0.135	0.037	0.045
	(0.152)	(0.144)	(0.103)	(0.102)
Unmarried, Cohabiting (d)	0.037	0.014	-0.021	-0.031
	(0.058)	(0.049)	(0.045)	(0.038)
Parent of at least one child (d)	-0.031	-0.033	-0.054	-0.044
	(0.040)	(0.039)	(0.033)	(0.033)
Less than high school (mother) (d)	0.056	0.056	-0.035	-0.052
	(0.098)	(0.096)	(0.064)	(0.045)
Some college (mother) (d)	-0.036	-0.030	0.075	0.031
	(0.101)	(0.101)	(0.193)	(0.141)
College (mother) (d)	0.031	0.026	0.008	-0.022
	(0.055)	(0.052)	(0.049)	(0.037)
Post-graduate (mother) (d)	-0.075**	-0.085***	0.044	0.022
	(0.036)	(0.029)	(0.073)	(0.062)
Less than high school (father) (d)	0.070	0.073	0.044	-0.010
	(0.088)	(0.092)	(0.085)	(0.061)
Some college (father) (d)	-0.029	-0.003	-0.003	0.005
	(0.050)	(0.057)	(0.056)	(0.056)
College (father) (d)	-0.020	0.007	0.073	0.074
	(0.050)	(0.054)	(0.066)	(0.059)
Post-graduate (father) (d)	-0.009	0.000	-0.095***	-0.086***
	(0.059)	(0.063)	(0.026)	(0.026)
Inverse Mills Ratio	0.059	-0.024	-0.441	-0.404
	(0.282)	(0.276)	(0.317)	(0.279)
Pseudo R-squared	0.18	0.20	0.28	0.29
N	518	518	517	517
Peak of age-risk parabola	29.45	29.00	53.03	54.13
Mean of dependent variable	0.14	0.14	0.15	0.15

The dependent variable is dichotomous equalling 1 for the outcomes listed in the column headers: whether the respondent had ever been assaulted by a client and whether the client had ever solicited a potential client from a public location such as a street, alley, parking lot, garage, highway, etc. Heteroskedastic robust standard errors in parenthesis. Coefficients have been transformed into marginal effects for ease of interpretation. SASP inverse probability weights are used in estimation, and all models include Census division fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 14** Estimated marginal effect of bodyweight on the managing new client uncertainty, SASP, probit

	Requires references	Googles new clients	Requires ID	Calls work	Uses gut instincts	Background check
Body Mass Index	0.001 (0.005)	0.006 (0.005)	0.001 (0.005)	-0.000 (0.001)	0.000 (0.000)	-0.009*** (0.003)
BMI < 18.5 (d)	-0.024 (0.124)	-0.089 (0.122)	0.015 (0.097)	0.132 (0.086)	0.035 (0.034)	-0.162*** (0.028)
25 ≤ BMI < 29.9 (d)	-0.044 (0.085)	-0.027 (0.084)	0.055 (0.078)	0.000 (0.022)	-0.000 (0.000)	-0.079** (0.035)
BMI ≥ 30 (d)	0.076 (0.092)	0.193** (0.079)	0.135 (0.097)	0.009 (0.025)	0.010 (0.014)	-0.094*** (0.036)
Age	0.004 (0.039)	0.010 (0.037)	0.011 (0.034)	0.006 (0.009)	-0.000 (0.000)	-0.003 (0.021)
Age-squared	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Years since began prostitution	-0.002 (0.005)	-0.006 (0.005)	0.003 (0.005)	-0.004** (0.002)	-0.000 (0.000)	-0.002 (0.003)
Black (d)	-0.202 (0.338)	0.192 (0.271)	-0.194 (0.184)	-0.023 (0.037)	-0.000 (0.001)	-0.120 (0.084)
Other Race (d)	0.110 (0.126)	0.098 (0.130)	0.179 (0.141)	-0.035*** (0.011)	0.185 (0.135)	0.000 (0.085)
Less than high school (d)	0.200 (0.160)	0.356*** (0.106)	0.406* (0.209)	(0.143)	(0.206)	0.000 (0.343)
Some college (d)	0.444*** (0.132)	0.440*** (0.151)	0.226 (0.151)	0.136* (0.081)	0.084 (0.110)	0.209 (0.183)
College grad (d)	0.489*** (0.107)	0.482*** (0.109)	0.300** (0.123)	0.188* (0.101)	0.007 (0.000)	0.312*** (0.111)
Married, Cohabiting (d)	0.102 (0.110)	0.107 (0.089)	0.116 (0.115)	-0.024* (0.014)	0.085 (0.097)	0.092 (0.098)
Divorced (d)	0.203** (0.080)	0.199** (0.080)	0.141 (0.095)	-0.041*** (0.013)	0.000 (0.000)	0.169* (0.089)
Separated (d)	0.178 (0.121)	0.181 (0.120)	-0.115 (0.103)	-0.033*** (0.010)	0.001 (0.003)	0.087 (0.139)
Unmarried, Cohabiting (d)	-0.119 (0.110)	-0.123 (0.109)	-0.059 (0.095)	-0.042*** (0.014)	-0.000 (0.000)	-0.046 (0.051)
Parent of at least one child (d)	-0.117 (0.088)	-0.114 (0.087)	0.084 (0.078)	0.007 (0.019)	-0.000 (0.000)	-0.002 (0.047)
Less than high school (mother) (d)	0.196 (0.131)	0.192 (0.131)	0.068 (0.151)	0.067 (0.079)	0.096 (0.133)	0.081 (0.107)
Some college (mother) (d)	0.279 (0.192)	0.067 (0.248)	0.281 (0.253)	-0.020 (0.041)	0.498 (0.244)	0.021 (0.152)
College (mother) (d)	0.055 (0.091)	0.057 (0.089)	0.179** (0.087)	-0.035* (0.018)	-0.001 (0.002)	0.150** (0.076)
Post-graduate (mother) (d)	0.131 (0.105)	0.133 (0.104)	0.097 (0.107)	-0.023* (0.014)	0.284** (0.117)	0.286** (0.118)
Less than high school (father) (d)	-0.049 (0.136)	-0.051 (0.136)	0.039 (0.116)	-0.038*** (0.012)	0.006 (0.000)	0.265** (0.117)
Some college (father) (d)	0.053 (0.104)	0.058 (0.103)	0.067 (0.102)	-0.017 (0.018)	0.001 (0.003)	-0.010 (0.067)
College (father) (d)	-0.113 (0.095)	-0.106 (0.094)	-0.044 (0.074)	-0.000 (0.024)	0.000 (0.000)	-0.061 (0.050)
Post-graduate (father) (d)	-0.133 (0.131)	-0.133 (0.132)	0.006 (0.111)	0.079 (0.051)	-0.000 (0.000)	-0.102** (0.047)
Inverse Mills Ratio	-0.896 (0.623)	-0.916 (0.618)	-0.341 (0.568)	-0.099 (0.138)	-0.003 (0.005)	-0.208 (0.348)
Pseudo R-squared	0.18	0.13	0.16	0.33	0.60	0.23
N	532	532	532	501	368	532
Peak of age-security parabola	8.39	3.78	20.99	48.73	39.30	157.54
Mean of dependent variable	0.60	0.57	0.32	0.11	0.03	0.21

The dependent variable is dichotomous equaling 1 for each of the column header outcomes: whether the respondent requires references before meeting a new client, whether they use search engines to learn about a new client, whether they require new clients to provide verifiable identification (e.g., driver's license), and whether they perform background checks on new clients beforehand. Heteroskedastic robust standard errors in parenthesis. Coefficients have been transformed into marginal effects for ease of interpretation. SASP inverse probability weights are used in estimation, and all models include Census division fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 15** Estimation of log of total payment, SASP, fixed effects estimator

Depvar: ln(payment)	Models separated according to:				
	Combined	Underweight	Normal	Overweight	Obese
Ln(length)	0.547*** (0.038)	0.519*** (0.090)	0.576*** (0.047)	0.456*** (0.069)	0.536*** (0.056)
Age of Client	0.003 (0.007)	-0.015 (0.039)	0.023** (0.011)	-0.022** (0.011)	-0.010 (0.008)
Age of Client Squared	0.000 (0.000)	0.000 (0.000)	-0.000* (0.000)	0.000* (0.000)	0.000 (0.000)
Client was a Regular	-0.015 (0.021)	-0.009 (0.091)	-0.047** (0.021)	0.036 (0.051)	-0.074 (0.052)
Asian Client	0.092** (0.039)	0.076 (0.135)	0.131** (0.059)	-0.038 (0.107)	-0.062 (0.061)
Black Client	0.076 (0.046)	0.180 (0.149)	0.012 (0.050)	-0.002 (0.082)	0.087 (0.078)
Hispanic Client	0.038 (0.087)	-0.159* (0.092)	0.064 (0.102)	0.097** (0.049)	-0.097 (0.175)
Other Ethnicity Client	0.085 (0.062)	-1.157*** (0.182)	0.208*** (0.077)	-0.018 (0.071)	0.123 (0.095)
Second provider present	0.065 (0.063)	0.390*** (0.118)	-0.006 (0.067)	0.321 (0.227)	-0.138 (0.092)
Met Client in Hotel	0.013 (0.038)	0.091 (0.088)	0.005 (0.052)	0.109* (0.060)	-0.074 (0.065)
Out-of-state travel for assignation	-0.043 (0.113)	-0.073 (0.107)	-0.108 (0.142)	0.086 (0.086)	0.174 (0.171)
Gave Client a Massage	-0.011 (0.047)	0.029 (0.101)	-0.008 (0.058)	0.090 (0.065)	-0.032 (0.030)
Cunnilingus	-0.008 (0.035)	-0.080 (0.091)	0.028 (0.042)	0.068 (0.052)	-0.067 (0.063)
Vaginal Sex with Condom	0.081* (0.042)	0.078 (0.115)	0.130* (0.070)	-0.006 (0.072)	0.072 (0.052)
Vaginal Sex without Condom	0.091 (0.162)	-0.377 (0.382)	0.233 (0.238)	-0.226 (0.137)	0.059 (0.091)
Fellatio with Condom	0.108** (0.052)	-0.105 (0.064)	0.140** (0.070)	0.158** (0.077)	0.033 (0.130)
Fellatio without Condom	0.066 (0.047)	0.007 (0.098)	0.052 (0.072)	0.108 (0.066)	0.026 (0.118)
Anal Sex with Condom	0.088** (0.036)	0.054 (0.093)	0.081 (0.049)	-0.026 (0.076)	-0.107** (0.053)
Anal Sex without Condom	0.181 (0.205)	0.244 (0.237)	0.274 (0.172)	0.488*** (0.157)	0.424** (0.162)
First introduction by word of mouth	-0.194 (0.128)	-0.114 (0.169)	-0.033 (0.184)	-0.268 (0.265)	
First introduction by telephone	0.039 (0.086)	0.156 (0.301)	0.137 (0.137)	-0.063 (0.168)	-0.204** (0.101)
First introduction through referral	-0.025 (0.075)		0.008 (0.127)	-0.024 (0.199)	-0.237* (0.127)
First introduction face to face	-0.051 (0.182)	0.589 (0.408)	0.097 (0.149)	-0.417 (0.255)	0.911*** (0.062)
First introduced to client via email/internet	0.017 (0.076)	0.431 (0.280)	0.075 (0.120)	0.013 (0.182)	-0.211** (0.085)
R-squared	0.58	0.69	0.61	0.63	0.73
N	1989	99	1226	342	228
Vertex of age-payment parabola	-552.09	40.84	54.82	45.26	41.23
Mean of dependent variable	5.76	5.87	5.85	5.58	5.42

We estimated the effect of transaction characteristics on the natural log of hourly adjusted prices for each group of women in our sample: those with BMI classified as underweight, normal, overweight and obese. We used the panel fixed effects estimator controlling for provider fixed effects, and corrected the standard errors for heteroskedasticity and clustering within a particular seller. SASP inverse probability weights are used in estimation, and all models include Census division fixed effects, as well as state-of-transaction fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01



**Table 16** Estimation of wage premia with obesity/overweight interactions, SASP, fixed effects estimator

Depvar: ln(payment)	Obesity interactions				Overweight interactions			
	1	2	3	4	1	2	3	4
Obese x vaginal sex with condom	-0.093 (0.083)							
Obese x vaginal sex without condom		0.129 (0.201)						
Obese x anal sex with condom			-0.221*** (0.066)					
Obese x anal sex without condom				0.587*** (0.202)				
Overweight x vaginal sex with condom					-0.069 (0.091)			
Overweight x vaginal sex without condom						-0.197 (0.186)		
Overweight x anal sex with condom							-0.050 (0.082)	
Overweight x anal sex without condom								0.157 (0.185)
Vaginal Sex with Condom	0.096* (0.054)	0.078* (0.045)	0.085* (0.045)	0.081* (0.044)	0.090* (0.049)	0.079* (0.045)	0.085* (0.045)	0.081* (0.044)
Vaginal Sex without Condom	0.131 (0.177)	0.111 (0.178)	0.127 (0.171)	0.136 (0.166)	0.123 (0.172)	0.136 (0.184)	0.126 (0.171)	0.139 (0.166)
Anal Sex with Condom	0.088** (0.037)	0.091** (0.036)	0.100*** (0.038)	0.089** (0.037)	0.092*** (0.036)	0.091** (0.036)	0.103** (0.042)	0.092** (0.037)
Anal Sex without Condom	0.159 (0.212)	0.157 (0.205)	0.153 (0.207)	0.090 (0.217)	0.170 (0.210)	0.159 (0.216)	0.154 (0.208)	0.150 (0.208)
R-squared	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58
N	1947	1947	1947	1947	1947	1947	1947	1947
Mean of dependent variable	5.76	5.76	5.76	5.76	5.76	5.76	5.76	5.76

We estimated the effect of the interaction of provider obesity/overweight with vaginal and anal receptive intercourse with and without a condom on the natural log of hourly adjusted prices for each group of women in our sample. We used the panel fixed effects estimator controlling for provider fixed effects, and corrected the standard errors for heteroskedasticity and clustered by seller. Because we are interacting the provider's body type with the sex act, and body type does not vary over the transactions, the obesity dummy variable is dropped, but the interaction is not. We control for but do not report the other covariates for the sake of brevity. SASP inverse probability weights are used in estimation, and all models include Census division fixed effects, as well as state-of-transaction fixed effects (available upon request). \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

## References

- Averett, Susan, Hope Corman and Nancy Reichman. 2012. “Effects of Overweight on Risky Sexual Behavior of Adolescent Girls.” *Economic Inquiry* Forthcoming.
- Averett, Susan and Sanders Korenman. 1996. “The Economic Reality of the Beauty Myth.” *Journal of Human Resources* 31(2):304–330.
- Cawley, John. 2004. “The Impact of Obesity on Wages.” *Journal of Human Resources* 39(2):452–474.
- Cawley, John, Kara Joyner and Jeffery Sobal. 2006. “Size Matters: The Influence of Adolescents’ Weight and Height on Dating and Sex.” *Rationality and Society* 18(1):67–94.
- Cunningham, Scott and Todd D. Kendall. 2010. “Moonlighting: Skill Premia in Commercialized Sex Markets.” Unpublished Manuscript.
- Cunningham, Scott and Todd D. Kendall. 2011a. “Men in Transit and Prostitution: Using Political Conventions as a Natural Experiment.” *The B.E. Journal of Economic Analysis and Policy* 11(1).
- Cunningham, Scott and Todd D. Kendall. 2011b. “Prostitution 2.0: The Changing Face of Sex Work.” *Journal of Urban Economics* 69:273–287.
- Cunningham, Scott and Todd Kendall. 2011c. *Handbook on Family Law and Economics*. Edward Elgar chapter “Prostitution, Technology, and the Law: New Data and Directions”.
- Edlund, Lena and Evelyn Korn. 2002. “A Theory of Prostitution.” *Journal of Political Economy* 110(1):181–214.
- Fletcher, Jason. 2009. “Beauty vs. Brains: Early Labor Market Outcomes of High School Graduates.” *Economics Letters* 105:321–325.

- Gale, David and Lloyd Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly* 69:9–15.
- Gambetta, Diego. 2010. *Codes of the Underworld*. Princeton University Press.
- Gertler, Paul, Manisha Shah and Stefano M. Bertozzi. 2005. "Risky Business: The Market for Unprotected Commercial Sex." *Journal of Political Economy* 113(3):518–550.
- Guista, Marina Della, Maria Laura Di Tommaso and Steinar Strom. 2009. "Who's Watching? The Market for Prostitution Services." *Journal of Population Economics* 22(2):501–516.
- Hamermesh, Daniel S. and Jeff E. Biddle. 1994. "Beauty and the Labor Market." *American Economic Review* 84(5):1174–1194.
- Hawthorne, Nathaniel. 1850. *The Scarlet Letter*. Ticknor, Reed and Fields.
- Laumann, Edward O., J. Gagnon, Robert T. Michael and S. Michaels. 1994. *The Social Organization of Sexuality*. University of Chicago Press.
- Logan, Trevon D. 2010. "Personal Characteristics, Sexual Behaviors, and Male Sex Work: A Quantitative Approach." *American Sociological Review* 75(5):679–704.
- Logan, Trevon and Manisha Shah. 2009. "Face Value: Information and Signaling in an Illegal Market." NBER Working Paper No. 14841.
- Mocan, Naci and Erdal Tekin. 2010. "Ugly Criminals." *Review of Economics and Statistics* 92(1):15–30.
- Mukhopadhyay, Sankar. 2008. "Do Women Value Marriage More? The Effect of Obesity on Cohabitation and Marriage in the USA." *Review of Economics of the Household* 6:111–126.
- Rao, Vijayendra, Indrani Gupta, Michael Lokshin and Smarajit Jana. 2003. "Sex Workers and the Cost of Safe Sex: The Compensating Differential for Condom Use among Calcutta Prostitutes." *Journal of Development Economics* 71:583–603.

- Robinson, Jonathan and Ethan Yeh. 2011a. "Risk-Coping through Sexual Networks: Evidence from Client Transfers in Kenya." *Journal of Human Resources* Forthcoming.
- Robinson, Jonathan and Ethan Yeh. 2011b. "Transactional Sex as a Response to Risk in Western Kenya." *American Economic Journal: Applied Economics* 3(1):35–64.
- Roth, Alvin E. 2008. "Deferred ACceptance Algorithms: History, Theory, Practice, and Open Questions." *International Journal of Game Theory* 36:537–569.
- Roth, Alvin E. and Marilda A. Oliveira Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monographs Cambridge University Press.
- Sabia, Joseph J. 2007. "The Effect of Body Weight on Adolescent Academic Performance." *Southern Economic Journal* 73(4):871–900.
- Sabia, Joseph J. and Daniel I. Rees. 2011. "The Effect of Body Weight on Adolescent Sexual Activity." *Health Economics* 20:1330–1348.
- Salop, Steven C. 1979. "Monopolist Competition with Outside Goods." *Bell Journal of Economics* 10:141–156.
- Shah, Manisha and Raj Arunachalam. 2008. "Prostitutes and Brides?" *American Economic Review Papers and Proceedings* 98(2):516–522.
- Weitzer, Ronald. 2005. "New Directions in Research on Prostitution." *Crime, Law & Social Change* 43:211–235.

## A TER Data

TheEroticReview is a “customer review” website devoted exclusively to prostitution reviews and functions much like Yelp.com or RateMyProfessor.com. Following assignation with a prostitute, a client with membership at TER may submit an online review. The review form in question has static fields related to detailed information on the prostitute’s physical characteristics, prices, services offered, as well as ratings (on a ten-point scale) of her overall appearance and “performance”. All workers reviewed on TER must have an internet “presence” of some form or another otherwise the review is not maintained in the database. For instance, this would require either the worker to have an email address, an online advertisement, or a personal website. Through selection both on the kinds of clients who use the site, as well as through requirements such as these, TER is a convenience sampling of the emergent “internet-facilitated” prostitution sector in the United States. The reviewing form demands an internet contact, such as email if available, and a telephone number for the worker. In addition, reviewers are asked to provide a detailed free-form narrative of their meeting with the encounter.

Over most of the history of TER, access to the information from these reviews, including contact information, was available to anyone with internet access for free. The exceptions are information on prices and specific services offered, and the more detailed parts of the narratives. Site users who submit two usable reviews in a month receive free access to the additional information; alternatively, users may purchase access for a fee (as we did in order to collect the data). In Summer 2009, the site began restricting more information to members.

Based on extensive telephone and email interviews with sex workers, as well as published writings about Internet prostitution (Brooks 2008, 2009), we believe TER reviews are extremely important mechanisms used to establish worker reputation as well aid in the matching process. Positive reviews are valuable resources, and workers exert substantial effort to maintain positive reviews on the site.<sup>20</sup>

---

<sup>20</sup>Feedback from bad reviews is rapid and may result in decreased earnings due to a slowdown in business. One interviewee remarked that “A lot of girls use TER for the ratings. I have dozens of pages of reviews [each page shows 10 reviews] on there. It took me years to get those good ratings, and I make good money now because of it, but I had to work my butt off for them.”

While TER reviews create reputations that are useful for prostitutes to generate business and revenue, the reviewing mechanism can be captured by clients for personal gain as well. For example, one interviewee shared a story about clients threatening to leave poor reviews as a credible threat to receive price or service concessions from sex workers. In a high-profile and well-publicized case, Dave Elms, the CEO and founder of TER, was arrested in 2008 on weapons and narcotics violations. The police investigation also turned up evidence that Elms had been extracting sexual favors from prostitutes reviewed on his site in exchange for removing damaging reviews about them (Richtel, 2008).

Other than through this type of activity, there are two major ways TER review data could potentially be corrupted. First, users could submit inauthentic reviews in order to gain access to the sites restricted price and narrative details. Such activity is likely to be rare since all other information, including contact and website information is available for free to all users, and price information can usually be determined independently through these means. Moreover, the review form is extremely detailed and takes several minutes to fill out, even with falsified information. Finally, users do not gain access to the price and narrative information until their reviews have been checked by TER staff, which takes several days, so impatient users attempting to gain access to restricted information would likely be dissuaded.

Second, sex workers may attempt to “review” themselves in order to appear to have more or better reviews than they actually do. In order to combat this type of fraud, the site allows users to “click through” a reviewers screen name to see all other reviews supplied by that client. It is common for the users of these review sites to inspect, not only the posted “field report”, but also the user leaving the report, as individual prostitutes with whom we spoke stated that they will often check the name of users leaving reviews before meeting with them for the first time, to look for irregular patterns that would suggest the person is a police officer. In like manner, users may discount reviews from persons who have not reviewed many other workers (as would likely be the case with this type of fraud).

A potential limitation, but also a strength, of the data is its temporal nature. Characteristics, services, and prices are posted on the site based on the initial reviewers observa-

tions. Additional users may review the same worker, and the appearance and performance ratings, as well as the user-supplied narratives will be grouped together on the site, but it appears that the original characteristics, services, and prices are generally not updated over time. Thus, the characteristics in the data are best thought of as representing “new entrants” to the site at a particular date. This fact about the data means there is relatively little within-provider temporal variation for researchers to exploit, but by the same token, it also allows researchers to observe changes over time in the characteristics and prices of sex workers appearing in the market at different dates, which can be matched to temporal location-specific economic and social conditions or particular events.

Another possible limitation is the potential for the same worker to be reviewed multiple times under different names. In general, TER appears to match new incoming reviews with individuals already reviewed on the site by telephone number, website URL, and email address, and the fact that many workers have scores, if not hundreds, or reviews, indicates that the matching process appears to work reasonably well. However, we have come across isolated cases of individuals who appear, from a comparison of photographs, to be the same person, but who are listed as two different people, perhaps because they changed their contact information. To the extent that such match failures were a random sample of all individuals reviewed, most analyses would be little affected; however, researchers should consider the potential effects on their results if workers who change contact information frequently are systematically different from others.

## **B SASP Survey Data**

Administering surveys to individuals involved in an illegal activity presents a number of problems to the researcher, besides the general concerns associated with survey data generally, all of which increase the cost of collecting such data. Additional precautions must also be taken to maintain ethical standards for research; a close relationship with the relevant institutional review board is a necessity.

First and foremost, the researcher must satisfy participants that (s)he is not a law enforcement officer, or an agent of a taxation authority , and that survey responses are

unlikely to be subpoenaed by a court and matched to a participant's true identity. Thus, anonymity is paramount. For SASP, potential respondents received an invitation to take the survey by email, which included a random string of characters and numbers generated by a third party inaccessible to the researchers (the information librarian at Baylor University). When surveys were returned, only the random string was observable by the researchers, not the email address of the respondent. We also allowed participants to take the survey by telephone with one of us or our research assistants if they felt uncomfortable responding electronically.

In the email used to invite survey responses, the anonymity of the survey was repeatedly emphasized. As additional signals that the surveyors were authentic academic researchers, the survey was hosted on Baylor University servers, and a website was posted with answers to frequently asked questions along with links to our personal websites, curricula vitae, institutional review board exemption letter, and research manuscripts. The office telephone number for one of us (Cunningham) was included in the email with an invitation to call for answers to any questions. Many did so, indicating that participants found this to be a useful means of verifying our authenticity.

A second difficulty associated with surveying sex workers involves controlling the flow of information during the survey period. Sex workers, especially those operating through the Internet, have developed substantial communications networks online, including private chat rooms and posting boards, as well as simple word-of-mouth links. During the period of implementation, we learned that SASP was a frequent topic of discussion through these channels, and there were apparently attempts to dissuade workers from responding to the survey. Since, as researchers, we did not have access to most of these private sites, we (and our assistants) engaged in continuous efforts to encourage responses, including, where possible, contacting our antagonists personally in an attempt to allay their concerns, and asking workers who were friendly towards us to post positive comments about the survey on sites where sex workers congregate. These efforts involved a substantial amount of time and effort, and we believe the survey likely would have been a complete failure otherwise.

Our survey design likely undersamples from several important subgroups of workers.



First, outdoor workers such as streetwalkers are unlikely to advertise online, and, we believe, are unlikely to be reviewed on TER. Second, workers employed in escort agencies or brothels frequently do not have personal email addresses listed either on TER or Eros.com; commonly, only the agency or brothel managers email is available. It seems likely that many of these workers were not reached. Thirdly, cases in which personal assistants or pimps are the primary contact for a worker are also unlikely to have been forwarded to the worker herself. Finally, since the majority of our contacts for the survey were workers reviewed by clients, we believe our population may miss some very high-priced workers, especially those who operate entirely through personal referrals.

After collecting all available contact information from TER and Eros.com, the list of potential survey participants was organized by city. We then randomly selected 4 or 5 cities per month, and attempted to contact individuals in those cities four times during that month. In addition to emailing participants with a link to the survey, we also attempted to contact a random selection of workers by telephone to encourage them to take the survey. We put the SASP survey in the field in August, 2008 and wrapped up data collection in early June, 2009.

The actual survey instrument (if completed online, and not over the telephone) was distributed and published, and the responses collected and organized, using SNAP 9.2 software, and included approximately 267 questions. Figure A1 shows the welcome screen participants saw when taking the survey, and the actual questions asked may be found in Table A1. Based on timestamps associated with participant answers, respondents took approximately 25 minutes to complete the survey.

Among the original 26,189 emails sent, 13,333 emails were successfully delivered. The high number of “bounce-backs” is unsurprising, given the fact that TER data stretches back to 1998, and many workers active in earlier years may have left the industry or changed contact information (similarly, imagine sending letters to all businesses listed in the last 10 years of telephone directories for a city). Consistent with this hypothesis, Figure A2 shows the number of emails collected from TER by the year of the workers first review, and the share of those emails that were undelivered. For workers first reviewed in 1999, nearly 90 percent of the emails listed were inoperable, while less than 20 percent of

emails sent to workers reviewed in 2009 were rejected.

While 13,333 emails were successfully delivered to a permanent email account, some share of these accounts likely remain open, even while the individual who once used them no longer checks the account regularly. Thus, this number represents an upper bound on the pool of potential participants, and the real response rate is likely much higher as a result. Nevertheless, between August, 2008 until June, 2009, 685 respondents answered our request to take the survey, giving us a lower bound response rate of 5.14 percent.

In order to extrapolate SASP responses to the population of TER-reviewed workers, we created probability weights for each respondent. Thus for instance, there are 1,155 White workers between ages 31 and 35 reviewed on TER, which is 11 percent of all TER-reviewed workers. Likewise, there were 99 White SASP respondents aged 31-35, which is 15 percent of all SASP respondents. The inverse probability of appearing in our sample is therefore 0.72 ( $= 0.11/0.15$ ) for Whites aged 31-35.

In order to gauge the reasonableness of our results using this methodology, we compared population-weighted means in SASP against those in a smaller survey by Church, et al. (2001). The latter survey involved interviews of 240 female prostitutes in the United Kingdom, including 125 indoor workers. In Table A2, we provide sample means for SASP and Church et al.s indoor sample on four questions which roughly overlap between the surveys. For current age, age at first entry into prostitution, and experience, mean answers between the surveys are similar. Notably, however, SASP respondents report considerably lower risk levels for client violence. Given the demographic similarities, this difference may suggest that SASP respondents are drawn from a pool of workers who take more care to screen out violent clients. We also compared the demographics of our sample with another survey, which focused on 100 internet-based sex workers (Milrod, 2008) and again found very similar age and race distributions. While a 5.14 percent lower-bound response rate is low by traditional survey standards, it is not low compared with other methodologies that use email solicitation to recruit respondents. Even surveys conducted by well-known surveyors on legal behavior report lower response rates using email opinion surveys. Zogby International, for instance, received a 14 percent response rate for a survey simply asking about individual opinions about public polices having

nothing to do with illegal behavior. As described in Butorivic and Klein (2010), “[t]he survey was administered by Zogby International by usual procedure. Because the panel is large it is difficult to maintain currency of all the email addresses in it. Zogby estimates a response rate of 14 percent and a completion rate of 10 percent”

In addition, it is unclear whether our response rate is low in comparison to other surveys of illegal or socially sanctioned behavior. For instance, Jones and Forrest (1992) found significant underreporting of the incidence of abortion, as did Allen (2007) with respect to rape. This is not to say that low response rates are not problematic. But, given the lack of perfect data on this important but hidden population (eg, prostitutes) there is both reason to be both cautious about inference and willing to take incremental steps on data collection with these constraints in mind.