An Equilibrium Search Model of the Retail Cocaine Market and Drug Law Enforcement

Beth Freeborn†
University of Virginia

October 2003

Abstract

This paper develops and estimates an equilibrium search model of the retail cocaine market. Consumers value the characteristics of a package of cocaine (price, weight and purity level) and follow an optimal purchasing strategy to maximize the value of search. Firms face a multidimensional production choice. The model allows firms (cocaine dealers) to choose the quality (purity) and weight of an individual package when maximizing profit. Drug law enforcement plays a primary role in this model. Both the consumers and dealers in the cocaine market incorporate the legal penalties associated with purchasing and distributing cocaine when making the respective optimization decisions. The advantage of this model is that it allows for equilibrium effects of policy changes; a consumer reacts to adjustments in available package types and a dealer adjusts package type in response to a change in demand. The model is estimated using package type data from the Drug Enforcement Agency and county-level penalty data from the National Judicial Reporting Program and Uniform Crime Reporting program. The estimated model is used to analyze the effects of potential changes in the structure of legal penalties concerning cocaine.

* Work in progress. I would like to thank my advisors, John Pepper, Steven Stern and Leora Friedberg for their invaluable comments, suggestions and encouragement. Peter Reuter provided useful comments and discussion. I gratefully acknowledge the University of Virginia’s Bankard Fund for Political Economy and the Miller Center of Public Affairs for financial support. Opinions reflect those of the author and do not necessarily reflect those of the granting organizations.

† Mailing address: Department of Economics, University of Virginia; P.O. Box 400182; Charlottesville, VA 22904-4182. E-mail address: freeborn@virginia.edu.
1 Introduction

The federal government spends over $19 billion on drug use and enforcement programs annually (ONDCP, 2003). Most of the federal spending focuses on eliminating the source of illegal drugs and reducing crime related to the market for drugs.¹ The primary goal of programs that punish individual drug offenders is to decrease the sale and consumption of drugs. If strict enforcement increases the costs of suppliers, and thus price, consumption will fall (See Figure 1). Given the policy goal of reducing consumption, most drug-related empirical literature has assumed enforcement increases price and focused on estimating the magnitude of the change in consumption in response to a price increase.² Few empirical studies have analyzed the supply side of the market or modeled dealer behavior in response to enforcement.³ Although changes in price arise from enforcement on both consumers and suppliers, most research ignores the firm (drug dealer) side of the market and estimates the demand elasticity given the assumption that enforcement increases equilibrium price.

During the time period of increased enforcement, the price of cocaine dropped dramatically. Figure 2 shows how enforcement on both consumers and dealers in the cocaine market may cause a reduction in price if demand shifts down more than supply. Using the traditional supply and demand analysis, there is a clear decrease in the level of consumption in response to heavy enforcement; however the change in price is ambiguous. Price data alone does not provide enough information to infer the impact of enforcement on one side of the market. An equilibrium model is necessary to provide a complete description of how enforcement affects both consumers and dealers in the cocaine market.

The goal of this research is to gain a better understanding of the underlying market structure for illicit drugs; I focus on cocaine because there is data available on packages. I

¹State government expenditures on drug law enforcement are more difficult to quantify. State spending for all corrections in fiscal year 2001 totaled $38.1 billion (NASBO, 2001), and drug offenders represent 20% of the state prison population. Note that this figure does not include non-corrections expenditures nor expenditures made at the local (municipality) level.
³Two exceptions that analyze supply-side factors are Yuan and Caulkins (1998) and Miron (2001). Yuan and Caulkins (1998) find no causality between drug seizures (a proxy for enforcement) and cocaine price. Miron (2001) uses data on legal goods to estimate the markup of cocaine that is due purely to prohibition. He finds cocaine prices are 2.5-5 times higher than the price would be if cocaine were legal and subject to taxation and regulation.
construct a structural model of consumer-dealer interactions and estimate the parameters of the model using price data from the Drug Enforcement Agency (DEA), supplemented with enforcement and demand data. To model the equilibrium, I apply an empirical search model to the market for retail cocaine. Equilibrium search models are most often applied to markets with imperfect information where it takes time to find an agent with the desired characteristics. Generally, individuals maximize personal wealth and firms maximize profit. For example, workers search for a wage offer while firms choose the wages given production technology or consumers search for the lowest price of a good as firms set price to maximize profit. Many theoretical models have been developed that generate sensible endogenous price distributions for both the labor and product market. Due to the nature of available data, empirical implementation of equilibrium search models is generally limited to the labor market.\footnote{Eckstein and Wolpin (1990) are the first to estimate an equilibrium search model. The authors estimate the parameters of a search model using both duration of unemployment and accepted wage data. Although the model performs poorly, this first empirical analysis of an equilibrium search model had a large impact on future work. For a review of the empirical literature, see van den Berg (1999) or Canals and Stern (1998).}

Within the model, consumers determine a purchasing strategy using an optimal stopping rule. Each consumer assigns a value to every combination of price, purity and weight of cocaine package. The consumer enters the market and samples a package of cocaine. If the value of the offered package is too low given her personal preferences and beliefs about the available packages, the consumer does not purchase and seeks another offer. Otherwise, the consumer purchases the package and stops searching. Cocaine dealers consider the behavior of the consumer when choosing what type of package to produce. An important consideration in this market is that retail cocaine is vertically differentiated. Drug dealers choose not only how much to supply, but also the weight of a single package and level of quality (purity) to supply. A dealer may substitute between the two product dimensions of quality and quantity.\footnote{As an example, consider this passage from Williams (1989): “He (the dealer) also has considerable latitude when it comes to the quality of the product once it leaves the importers’ hands. One day he gives me a demonstration ... ‘This is fresh stuff’ he says, taking a few of the flaky particles and rubbing them between his forefinger and thumb. ‘If it’s a new customer, I might take thirty grams of flake and cut a little of it’ he explains. In other words, he might remove 30 grams of the uncut drug for his own profit or pleasure, and replace that with adulterant” (p.35).} Choice of purity on the firm side of the market has not been modeled previously, yet a great deal of purity adjustments occurs at the retail level.\footnote{National Drug Threat Assessment (2001).} Dealers choose package type
to maximize profit, where the cost of production includes the expected penalty associated with selling narcotics. While the lack of detailed data prevents me from modelling certain important characteristics of the drug market (e.g., addiction and reputation effects), this research is the first to directly model the interactions of consumers and dealers.

Consumer product search literature attempts to explain why price dispersion might exist in a homogenous good market, yet few of these models are empirically estimated. One example is Sorensen (2000) who must assume consumers search either exhaustively, or not at all, for prescription drugs. Without specifying an explicit model of search, he is able to link costly consumer search to price dispersion in retail prices at local pharmacies. Hortaçsu and Syverson (2002) use price and quantity data in the mutual fund market to estimate a model of search over differentiated products by investors. Often the only available data on consumer products are prices. Hong and Shum (2001) use price data and the equilibrium restrictions implied by a theoretical model to estimate search costs of consumers making purchases on-line.

This paper extends the equilibrium search literature into a product market where consumers value multiple characteristics and search to maximize the utility of consumption rather than personal wealth. I use the price and production choices of dealers to estimate a model of search for a differentiated product. The costs of search are observed in the market. Consumers vary in preferences and dealers vary in the costs associated with selling cocaine. The heterogeneity on both the consumer and firm side of the model allow for a non-degenerate price distribution.

Within the drug economics literature, this paper is most closely related to the work of Kuziemko and Levitt (2001) and Bushway, Caulkins and Reuter (2003) which analyze the role of law enforcement in the retail cocaine market. Kuziemko and Levitt estimate a reduced form regression model of the street level price of a pure gram of cocaine on certainty and severity of punishment. Bushway et. al. develop an empirical model of expected costs

---

7 van den Berg (1999) illustrates that the empirical search literature gives reliable estimates (i.e., matches the wage distribution well) only when both duration and wage data are jointly used. This fact may account for the lack of empirical search literature applied to consumer products; it is rare to find duration data on the search for a consumer good. However, if the econometrician is willing to place parametric assumptions on the distribution of search costs, an equilibrium model can be identified without duration data. Duration data is not necessary if the costs of search are observed.

8 These authors use the term 'enforcement' to define the level or severity of punishment drug offenders receive. I also use the term enforcement in reference to the punishment variables.
using a theoretical framework that implies an increase in enforcement leads to an increase in price. The authors estimate this expected cost model using county level enforcement variables and show a small but positive relationship between expected sentence term and retail price. Unlike the previous literature on drug markets, I model the supply and demand for retail cocaine jointly, allowing for equilibrium consequences of law enforcement.

The paper is organized as follows. Section 2 introduces the model. Section 3 describes the data necessary for estimating the empirical model and Section 4 presents some preliminary analysis of this data. Estimation strategies are discussed in Section 5 and Section 6 presents the results from estimation. Policy simulations are described in Section 7 and Section 8 concludes.

2 Economic Model

A consumer in the model maximizes the utility of consuming cocaine. She searches for packages of cocaine over a known distribution of package types and chooses whether or not to purchase using an optimal stopping rule.\(^9\) The market for available packages is defined at a county level.\(^10\) Given an offer of a particular package type, the benefit of search is the increase in utility from consuming a better package type. The cost of search is defined as a function of the county level legal penalty for possession of cocaine. A consumer may encounter an undercover law enforcement official posing as a dealer on her next sample and be arrested. The sample of the next package might also be observed by law enforcement, resulting in arrest. Each consumer determines a reservation value that maximizes the value of search. The reservation value splits the set of packages; she purchases any package that gives a value higher than her reservation value and does not purchase any package that has value less than the reservation value. The consumer stays in the market until she finds a package to purchase.

The cocaine dealer chooses the package type, defined by price, weight and purity level, in

\(^9\)It is not unreasonable to assume consumers have some beliefs about what types of packages are available. At one point, there existed a website that published average prices and weights of packages using voluntary consumer input. This website was located at http://www10.brinkster.com/brdrlne/.

\(^10\)The assumption of a county level market is due to data restrictions. If consumers cross county lines when looking to purchase cocaine, then defining the market as a county may be too limiting. However, if consumers do not leave their own neighborhood when looking to purchase cocaine, then county defines the market too widely.
order to maximize profit conditional on the behavior of consumers and other dealers. The
cost of operating in the retail market is a function of the probability of arrest and the penalty
associated with selling cocaine, both measured at the county level. Each dealer remains in
the market until he has sold all of his packages of cocaine.

The distribution of consumer reservation values is dependent upon the distribution of
profit-maximizing packages within the county. Each dealer maximizes profit conditional on
the optimizing behavior of consumers, resulting in a distribution of package types. Equilib-
rium is defined by a particular distribution of package types and the resulting distribution
of reservation values such that each dealer is maximizing his profit and each consumer is
maximizing her value of search.

The rest of this section describes the model in detail. Section 2.1 describes the consumer
side of the model; the dealer optimization problem is discussed in Section 2.2 and Section
2.3 briefly describes how an equilibrium to the model is found.

2.1 Consumer Behavior

Each county $l$ has a distribution of packages of cocaine $F_l(x)$ which is determined as a
solution to the set of conditions for dealers and consumers in that county. The consumer takes
the behavior of dealers as given and chooses a cocaine purchasing strategy that maximizes
the value of search.

2.1.1 Value of Consuming Cocaine

Each consumer chooses to consume cocaine or the outside option (to not consume co-
caine). The utility of the outside option is normalized to zero, and the utility of consuming
a package of cocaine $x$ is defined as $V = V(x, \theta)$ given preferences $\theta$. The vector $x = (p, q, s)$
defines a package of cocaine by the price $p$, weight $q$ and purity level $s$. When a consumer
is approached by a dealer offering a package of cocaine she learns the exact characteristics
(price, weight and purity) of the package. This is a necessary assumption of the model and
is common to search literature.\footnote{Within drug economics literature, there has not been a general consensus on how to appropriately treat purity. Caulkins (1994) argues that consumers have an expected level of purity that governs the price of a transaction and mean purity should be used to create prices. However, the widely used Abt Associates price series data (Rhodes, 1997) implicitly assumes that consumers know the purity at the time of the purchase. Most search literature assumes agents can perfectly observe characteristics when the good is}
A consumer knows her own value of $\theta$ when evaluating package $x$. The individual realization of $\theta$ is unobserved to the dealer and econometrician. The value of package $x$ to consumer $i$ is

$$V_i = \beta_i q^{\lambda_{qi}} s^{\lambda_{si}} - \alpha_ip.$$  

This functional form ensures the consumer assigns positive value only to the consumption of a package of cocaine with both positive weight and purity. The parameter vector representing preferences for characteristics of cocaine, $\theta = (\beta, \lambda_q, \lambda_s, \alpha)$ varies over consumers.\(^{12}\) These preferences are distributed $\Phi(\theta, \Omega)$, where $\theta$ represents the mean level of preferences and $\Omega$ is the covariance matrix of the distribution of preferences.\(^{13}\)

Every consumer values cocaine consumption according to equation (1), but not every consumer chooses to purchase cocaine. A consumer with a negative value of $\beta$ never receives positive utility from cocaine and chooses not to participate in the market.

The distribution of consumption values over package types for consumer $i$ in county $l$ is

$$G_{il}(v) = \Pr\{V(x, \theta) \leq v\} = \iiint_{\{x: V(x, \theta) \leq v\}} dF_i(x) .$$  

(2)

### 2.1.2 Expected Penalty for Possession

Each time a consumer searches for a cocaine package there is a fixed probability she may be arrested for possession of cocaine and incur the penalty. Thus, the expected penalty for possession is defined as:

$$c_l = \left[\Pr(A_l^P) \times (Possession\ Penalty)_l\right]$$

\(^{12}\)Ideally the preference parameters would vary over type of cocaine as well; $\theta$ would be indexed by $t$ where $t$ indicated which form of cocaine (crack cocaine versus powder cocaine). Given either consumer specific data on drug choices or dealer specific data on the drug types offered, I could estimate the covariance of preferences across drug forms. The available data does not allow for for identification of the covariance of preferences between powder cocaine, crack or any other drugs. Instead, I estimate the model separately for the different market forms. In general, not allowing for substitution into other narcotics could potentially lead to overestimating the preferences for the drug of study.

\(^{13}\)This is a standard representation of preferences for a consumer good. I do not model addiction explicitly, however the preferences for characteristics will capture the nature of addiction. Consumers who are addicted value consumption of cocaine differently than non-addicted consumers.
where \( \Pr(A^P_l) \) is the probability of being arrested for possession in county \( l \). The cost of search is defined as \( \delta c_l \) where the parameter \( \delta \) acts as a weighting coefficient in order to convert the cost of search defined as an expected penalty into the same terms as utility.

### 2.1.3 Reservation Value

Each consumer encounters a cocaine dealer who offers her a particular package of cocaine. Let \( V' \) be the value the consumer assigns to the current package. Given knowledge of the distribution \( F_l(x) \), the expected benefit of searching for a new offer is defined as

\[
H_i(V') = \int_{\{V \geq V'\}} [V - V'] dG_{il}(V)
\]

where \( H_i(V') \) represents consumer \( i \)'s expected gain in utility of finding a better package type given the current sample \( V' \).

The consumer follows an optimal stopping rule. If the cost of searching for another package is higher than the expected gain in utility from finding the next package, she buys the offered package and exits the market. If the expected benefit of sampling another package outweighs the cost, the consumer does not buy and looks for a new package type. The consumer continues to sample packages until the expected benefit from search is less than the cost of search.

The reservation value for consumer \( i \) with preferences \( \theta \) is defined to be \( R(\theta) \) such that

\[
H_i(R(\theta)) = \int_{\{V \geq R(\theta)\}} [V - R(\theta)] dG_{il}(V) = \delta c_l
\]

where the cost of search \( \delta c_l \) is the expected penalty for possession of cocaine in county \( l \). Each consumer \( i \) with preferences \( \theta \) has a reservation value defined where the cost of search equals the expected benefit of search.\(^{14}\) A consumer may find that the expected benefit of search given the distribution of packages does not outweigh the cost of search. This consumer is defined as participating in the market and choosing not to search until the expected benefit outweighs the cost of search.

Each consumer sets a reservation value \( R(\theta) \). The distribution of reservation values for all consumers is given by \( F_R \).

\(^{14}\)See the appendix for the proof of the existence of a reservation value.
2.1.4 Probability of Purchase

The consumer evaluates a package $x$ and chooses to purchase it based on the value of the package $V(x, \theta)$ relative to her reservation value $R(\theta)$. Prior to sampling an actual package offer, the consumer knows only her reservation value $R(\theta)$. The consumer purchases any package that gives her a value higher than her reservation value and rejects all packages returning a lower value.\(^{15}\)

The probability that consumer $i$ purchases $x$ is the probability that the value of $x$ is higher than her reservation value $R(\theta)$:

$$\Pr(Purchase) = \Pr\{V(x, \theta) \geq R(\theta)\} = 1 - G_u[R(\theta)].$$  \hspace{1cm} (4)

Each consumer purchases only one package at a time in order to avoid stricter penalties.\(^{16}\)

For consumer $i$, the quantity demanded of package $x$ is

$$d_i(x) = \begin{cases} 
1 & \text{if } V(x, \theta) \geq R(\theta) \\
0 & \text{if } V(x, \theta) < R(\theta)
\end{cases}. \hspace{1cm} (5)$$

This unit demand will have implications for the dealer behavior discussed in the next section.

2.2 Dealer Behavior

Each dealer knows the distribution of consumer preferences $\Phi(\theta)$ and the distribution of all packages offered to the consumers $F_l(x)$ (i.e., the choices of all other dealers in the market). The distribution of reservation values $F_R$ is inferred from the consumer preferences and available packages. The dealer takes the behavior of consumers and other dealers as given and chooses the profit-maximizing type of package to sell.

2.2.1 Number of Packages

Each dealer $j$ receives a stock of cocaine $K\varepsilon_j$ from a wholesale supplier. $K$ is the average pure-weight of the wholesale stock, and $\varepsilon_j$ is a dealer-specific non-negative error with mean-1

\(^{15}\)There are multiple combinations of price, purity and weight resulting in the same level of $V(x, \theta) = R(\theta)$. The reservation value for each consumer is a manifold in $p-q-s$ space dividing the space of packages into types that the consumer purchases and types the consumer does not purchase.

\(^{16}\)A consumer is assumed not to keep inventories of cocaine because holding large quantities of drugs causes police to infer an intent to sell (i.e., the consumer would be considered a dealer). Each consumer considers the value of consuming cocaine only in the current period. When the consumer needs to purchase another package, it signals the beginning of a new period and a new search process.
representing variation in the wholesale stock of cocaine. Each dealer pays a common fixed cost $T$ for the wholesale stock of cocaine. The marginal costs of production are assumed to be zero.$^{17}$

The number of packages $N_j$ the dealer sells is defined by the choice of weight $q_j$ and purity $s_j$:

$$N_j = \frac{K\varepsilon_j}{s_j \ast (q_j^{\omega_j})}.$$  \hspace{1cm} (6)

$\omega_j$ is a dealer specific non-negative error with mean-1 representing variation across dealers in their choice of substitution between purity $s$ and weight $q$ of an individual package.$^{18}$

### 2.2.2 Making a sale

The probability $P^j$ that dealer $j$ makes a sale is equal to the proportion of consumers in the market that will purchase his package $x_j$.

For every package $x_{j^*}$, there exists a set $S_{\theta^*}$ of preferences $\theta^*$ representing the preferences of the marginal consumers who choose to purchase $x_{j^*}$. Each element $\theta^*$ in the set $S_{\theta^*}$ is a marginal consumer that divides the set of all consumers into the set of consumers that purchase package $x_{j^*}$ and the set that do not purchase $x_{j^*}$.

Begin with a particular marginal consumer $\theta^*$ and let $\theta_m$ represent the $m^{th}$ element of $\theta$. Let $V(x_{j^*}, \theta)$ be increasing in $\theta_m$.\hspace{1cm}$^{19}$ A consumer with preferences $\theta$ such that one element $\theta_m$ is greater than $\theta^*_m$ and all other elements of $\theta$ are equal to the corresponding elements of $\theta^*$, chooses not to purchase the package. Her preferences require a package with ‘better’ characteristics than $x_{j^*}$. Any consumer with $\theta$ such that a single element $\theta_m$ is less than $\theta^*_m$ chooses to purchase the package $x_{j^*}$ (holding all other elements of $\theta$ constant at $\theta^*$).\hspace{1cm}$^{20}$

---

$^{17}$I assume MC = 0 for simplicity. I define the industry as one of sunk costs. Once a dealer chooses to be in the market and purchases wholesale cocaine, the costs of production are a function of whether or not he is able to sell the cocaine quickly and avoid arrest.

$^{18}$One could think of $\omega_j$ as the ‘skimming’ error. It is documented that street level dealers occasionally take some of the wholesale stock for personal use or give away some amount to friends at parties. This error allows for two different dealers who begin with the same initial stock $(K\varepsilon_j)$ to sell the same single weight package on the street. The difference in skimming amount caused by the dealer-specific error $\omega_j$ results in each dealer selling a package of a different purity level. $\omega_j$ is the elasticity of $N_j$ with respect to weight of package $q_j$.

$^{19}$The value of consuming cocaine, $V(x, \theta)$ is not increasing in all elements of $\theta$. I use $V$ increasing in $\theta_m$ (e.g., $\theta_m = \beta$) to aid in exposition of the concept.

$^{20}$Consider the simple example where consumers have identical preferences for all characteristics of cocaine other than price (identical $\beta, \lambda_q, \lambda_s$ but different values for $\alpha$). If consumer 1 with preferences $\alpha_1$ purchases package $j^*$, then consumer 2 with preferences $\alpha_2 < \alpha_1$ assigns a higher value to package $j^*$ than consumer 1. If $V_2(x_{j^*}) > R_2$ consumer 2 also purchases package $j^*$. There is some marginal consumer with preferences
This is true for all elements $\theta^*$ of the set $S_{\theta^*}$. For each package $x_{j^*}$ there are multiple values of $\beta, \lambda_q, \lambda_s$, and $\alpha$ such that the consumer is indifferent between purchasing $x_{j^*}$ and not purchasing. $S_{\theta^*}$ is a manifold in $\beta \times \lambda_q \times \lambda_s \times \alpha$ space that divides the set of all consumers. This is shown to be true by proving that $\Pr(V > R)$ is monotone in $\theta$. See the appendix for this proof.

Each consumer whose value of package $x_j$ is greater than her reservation value will purchase from dealer $j$ with certainty. Integrating over the set of consumers who will purchase $x_j$ gives the probability that dealer $j$ sells his package $x_j$:

$$P^j = \int 1 \{V(x_j, \theta) \geq R(\theta)\} \, d\Phi(\theta)$$

$$= \int_{\{\theta: V(x_j, \theta) > R(\theta)\}} d\Phi(\theta)$$

(7)

2.2.3 Number of Approaches

The stochastic number of approaches necessary to successfully sell $N_j$ packages is defined as $D_j$. The probability that dealer $j$ sells package $x_j$ is defined as $P^j$ in section 2.2.2. If there is some positive probability of failing to make a sale on any given approach $(1 - P^j)$, the total number of approaches $D_j$ is greater than the number of packages $N_j$.

$D_j$ is random with negative binomial distribution, $D_j \sim H_d(P^j, N_j)$ with

$$E(D_j) = \frac{N_j}{P^j}$$

(8)

2.2.4 Expected Profit

Each dealer chooses the weight, purity and price of a single package to maximize expected profit. A dealer is assumed to sell multiple packages of the same type. The number of packages dealer $j$ sells is defined to be $N_j$ and is determined by the choice of weight and purity. The number of consumers an individual dealer approaches is defined as $D_j$. The expected profit of dealer $j$ is given by:

$$E(\Pi_j) = \sum_{D_j} (1 - \Pr(A_j^S))^{D_j} \Pr(D_j) \ast p_j \ast N_j$$

$\alpha^*$ that divides the set of consumers into consumers who purchase package $j^*$ and consumers who choose not to purchase the particular package.

21 Recall the unit demand from Section 2.1.5

22 Given the limitations of the data, this assumption is necessary to identify the model.
\[-[1- \sum_{D_j} (1 - \Pr(A^S_j))^{D_j}[\Pr(D_j)]][\gamma W_l + \eta_j] - T. \quad (9)\]

The first term of the expected profit function represents total revenue if the dealer avoids arrest on all \(D_j\) approaches and sells \(N_j\) packages at a price \(p_j\); the second term is the penalty he pays if he is arrested on any single approach and the third term is the fixed cost of purchasing wholesale cocaine.

\(\Pr(A^S_l)\) is the probability of arrest for selling cocaine in county \(l\), \(\Pr(D_j)\) is the probability of making \(D_j\) approaches and \(\eta_j\) is a dealer-specific mean zero error term.\(^{23}\) \(W_l\) is a vector of the median prison sentence and median probation term in county \(l\) (both of which are transformed to account for discounting future periods) and \(\gamma\) is a weighting term on this penalty variable.\(^{24}\) \(T\) is the fixed cost of operating in the retail cocaine market.\(^{25}\)

The functional form of the expected profit given by equation (9) specifies that a dealer receives revenue only if he is not arrested on any of the approaches. If he is not arrested the dealer receives the same price on each package that he takes to the market and sells. The cost (i.e. penalty) is accrued only when he is arrested, in which case he does not earn any revenue.

The probability that the dealer is arrested on one approach and incurs the cost is increasing in the number of approaches. The fixed probability of arrest occurs with every approach made by the dealer rather than with the actual number of sales \(N_j\) (e.g., the next consumer may not purchase the package yet may be an undercover agent). The probability of a dealer receiving the penalty is \([1- \sum_{D_j} (1 - \Pr(A^S))^{D_j}[\Pr(D_j)]].\)

The probability of not being arrested on any approach \(\sum_{D_j} (1 - \Pr(A^S))^{D_j}[\Pr(D_j)]\) can be expressed as \(E(1 - \Pr(A^S))^{D_j}.\) Given the distribution of \(D_j\) and the probability of making a sale, the expected profit is thus:\(^{26}\)

\[
E[I_j(x, P^j(x))] = \frac{[P^j(1 - \Pr(A^S))]^{N_j}}{[1 - (1 - P^j)(1 - \Pr(A^S))]^{N_j}} * p_j * N_j
\]

\(^{23}\)There exists variation in the penalty a dealer receives once he is arrested. One dealer may have a special relationship with law enforcement, prosecutors, etc. causing (negative) variation (i.e., his expected penalty is lower than the mean). Another dealer may have a previous record, resulting in (positive) variation in the expected penalty.

\(^{24}\)See the appendix for a discussion of how the penalty lengths are transformed to account for discounting.

\(^{25}\)This fixed cost determines whether a dealer operates in the retail market. An individual does not choose to be a dealer if he makes negative profit. Instead, he pursues his next best option. This is a necessary component of the model for counterfactual analysis; if the penalty for sale of cocaine is increased some dealers will make negative profit and leave the market.

\(^{26}\)See the appendix for the derivation of the profit function.
\[-[1 - \frac{[P_j(1 - \Pr(A^S_j))]^{N_j}}{[1 - (1 - P_j)(1 - \Pr(A^S_j))]} \ast [\gamma W_i + \eta_j] - T. \quad (10)\]

A dealer chooses the profit-maximizing type of package to sell conditional on the distribution of consumer preferences \( \Phi(\theta) \) and his dealer specific errors, \( \varepsilon_j, \omega_j \) and \( \eta_j \). The probability of making a sale \( P^j \) is a function of the choice of dealer \( j \), the choices of all other dealers and the distribution of reservation values. Dealer \( j \) chooses

\[ x_j = \arg \max_x E [\Pi_j(x, P^j(x_j))]. \quad (11)\]

### 2.3 Equilibrium

Each consumer takes the cost of search \( c_l \) and her preferences \( \theta \) as given and sets a reservation value conditional on the distribution of package types available in the market. The set of acceptable packages for consumer \( i \) is the set of all packages that she would purchase:

\[ A_i = \{ x : V(x, \theta) \geq R(\theta) \}. \]

The set of packages that would be accepted by at least one consumer is the union of all consumers’ acceptable sets:

\[ A = \bigcup_{i=1}^{I} A_i \]

where \( i \) indexes the consumers in county \( l \).

The penalty for sale of cocaine \( W \), the distribution of dealer-specific errors \( F_e \) and the wholesale stock of cocaine \( K \) are exogenous. The dealer takes his draw of errors \( e_j \), the behavior of other dealers and the distribution of preferences of consumers \( \Phi \) as given. The dealer chooses his profit-maximizing type of package conditional on his probability of making a sale derived from the distribution of reservation values (see equation (11)).

At the equilibrium, no dealer chooses to produce a package that falls outside of the set \( A \) because there are no consumers who purchase that package. The set \( A \) defines the set of packages that are feasible in equilibrium, and the distribution \( F^E(x) \) defines the equilibrium distribution of packages.

**Definition 1** An equilibrium is defined by \( [F^E_R, F^E(x)] \) such that:

1. Each consumer maximizes utility by following a reservation value stopping rule taking the behavior of dealers as given. The distribution of optimal reservation values is \( F^E_R \).
2. Each dealer chooses production to maximize profit conditional on the behavior of consumers and the package type choice of all other dealers. The optimal distribution of packages is given by $F^E(x)$.

Equilibrium is defined as a distribution of reservation values $F_R^*$ and a distribution of package types $F(x)$ that are consistent with one another. Given the behavior of the dealer, the equilibrium distribution of reservation values $F_R^*$ generates profit-maximizing production decisions by each dealer, resulting in an equilibrium distribution of packages $F^*(x)$. This distribution of packages $F^*(x)$ generates the original distribution of reservation values $F_R^*$.

The reservation value $R(\theta)$ is determined implicitly by equation (3) for each consumer with preferences $\theta$. Given the distribution of $\theta$, there exists a distribution of reservation values $F_R$. $\Psi_R$ represents the function that maps the cross product of elements from the space of distributions of packages and the space of distributions of preferences into the space of distributions of reservation values. An element in the space of distributions of package types is $F(x)$ and an element from the space of reservation value distributions is $F_R$. The function is defined: $\Psi_R: F(x) \times \Phi \rightarrow F_R$.

The profit-maximizing choice of the dealer is conditional on his draw from the error distribution $F_e$ and satisfies equation (11) given the probability of making a sale $P_j$. The solution to the profit-maximizing problem for all dealers results in the distribution of packages $F(x)$. Let $\Psi_x$ represent the function that maps the cross product of elements from the space of distributions of reservation values and the space of dealer-specific error distributions $F_e$ into the space of distributions of package types $F(x)$: $\Psi_x: F_R \times F_e \rightarrow F(x)$.

Equilibrium is defined as a distribution of reservation values that generates itself: $\Psi_E: F_R \times \Phi \times F_e \rightarrow F_R$.

$$\Psi_E(F_R) = \Psi_R(\Psi_x(F_R, F_e), \Phi)$$

$$\Psi_E : F_R = \Psi_R(\Psi_x(F_R, F_e), \Phi)$$

Given the mapping $\Psi_E$, proof of the existence of the equilibrium follows from Schauder’s fixed point theorem.\footnote{See Appendix for necessary conditions for proof of the equilibrium.}
3 Data Description

In order to estimate the econometric model described above, I compiled data from a number of different sources. Specifically, I obtained data on the geographic distribution of package types and matched this to the legal penalty associated with sale and possession of cocaine in that particular county.

3.1 Purchase Data

The Drug Enforcement Administration (DEA) initiated a program to collect and manage data on drug purchases and seizures in the late 1970’s. The System to Retrieve Information from Drug Evidence (STRIDE) contains records of cocaine and heroin purchases and seizures in the United States. The data for this research is limited to cocaine purchases from January 1986 through December 2000. The purchase records include information regarding the drug’s form (crack versus powdered cocaine), weight, purity, price, location and date of purchase. The purchase cost is the price for the observed/seized package of cocaine, and weight is the total weight of the package in grams. The data includes 11866 observations spanning 185 counties (37 states) in the United States.

I define a retail package of cocaine as any package weighing less than three grams.\textsuperscript{28} Table 1 provides descriptive statistics of the STRIDE data used for this analysis. The average retail package weight is 0.67 grams with a minimum weight of a single package of 0.013 grams. The lowest purity level is 0.002 percent and the highest is 100 percent pure. The average purity level of a package in the retail market is 74 percent. Price for any retail package of cocaine ranges from $1 to $6000 and the average price in the retail market is $82.\textsuperscript{29}

The STRIDE data includes records of drug purchases made by DEA agents or informants in order to further criminal investigations and make arrests. Researchers using the STRIDE data argue that observed prices and package types must be relatively accurate. An unreasonable price would invoke suspicion on the part of the dealer and endanger the DEA agent. Although the observed packages fall within the distribution of market package types there

\textsuperscript{28}There is no general consensus on what weight threshold determines the retail cocaine market. Three grams is equivalent to roughly ten usage sessions. Abt Associates (1997), Yuan and Caulkins (1998), Bushway et.al (2003) and Kuziemko and Levitt (2001) use 1, 3.5, 28 and 140 grams, respectively to define the retail market.

\textsuperscript{29}The average price for a pure gram of cocaine is $272.
is concern that the STRIDE data does not represent a random sample of the distribution. For example, there may be variation in the criteria to begin an investigation over time and across offices (e.g., DEA agents in port cities may be oriented towards large purchases or seizures and less concerned with the retail market).\textsuperscript{30}

Given the criticism of the data, one may question the results and policy implications that arise from the structural model described above.\textsuperscript{31} The model is a new attempt at structuring the decisions of drug dealers and consumers at the retail level and evaluating how the agents’ decisions change in response to new policy. The policy implications that result from the model would be more informative if STRIDE represented a true random sample of prices from the illegal drug market. Although a more reliable source of price information would be welcome for estimation and counterfactual analysis, STRIDE is the most complete and widely used data set available. This exercise underscores the need for more accurate data so that structural models such as this one can be effectively used for policy evaluation.

3.2 Legal Penalty Data

Agents consider the magnitude of the penalty as well as the probability of incurring the penalty. The legal penalty data comes from two sources: the National Judicial Reporting Program (NJRP) and the Uniform Crime Reports (UCR). The prison and probation sentence length is collected from the NJRP. This data is used to construct the penalty term $W$ in the dealer profit function. The data used to calculate arrest rates is collected from the UCR. The arrest rate for possession is used in conjunction with the sentence length to construct the cost of search (expected penalty for possession) on the consumer side of the market. The arrest rate for sale is fundamental to determining the expected profit of the dealer.

3.2.1 Sentence Lengths: NJRP

The penalty for selling cocaine at the retail level may be a minimum jail or prison sentence, community service, probation or a monetary fine. In the model described above,

\textsuperscript{30}See Horowitz (2002).
\textsuperscript{31}The Committee on Data and Research for Policy on Illegal Drugs concluded that there is a pressing need to improve existing data and acquire more reliable drug price data. Until more reliable data is gathered, the Committee report states the nation will continue to be poorly informed about the price of illegal drugs, total expenditure on illegal drugs and the effectiveness of intervention techniques. See NRC (2001).
cocaine dealers consider the county level penalty when making their respective package type decision. The county level data on penalties is taken from the National Judicial Reporting Program (NJRP). NJRP is data collected every two years, surveying a nationwide sample of county felony trial courts in 344 counties. The NJRP contains detailed information on demographic characteristics of felons, conviction offenses, type of sentences and sentence lengths. The data is used to find the county level median sentence and probation/parole length for both possession and sale of cocaine. These terms are used to construct the penalty term \( W \) in the expected profit function.\(^{32}\) Table 1 shows the average probation and prison sentence for possession and sale of cocaine when matched with the STRIDE purchase data. On average, consumers receive longer probation terms than dealers. The average probation term for possession of cocaine is four times the average probation term for sale of cocaine. Dealers receive a longer prison sentence than consumers; the sentence for sale of cocaine is 148 percent the sentence for possession of cocaine.\(^{33}\)

Table 2 presents descriptive statistics of the NJRP data alone. Conditional on being arrested and found guilty of possession of cocaine, 87 percent of consumers are sentenced to a term between one month and five years. Relatively fewer dealers receive a similar sentence; seventy percent of sentences for sale of cocaine are between one month and five years. A dealer is more likely to receive a longer sentence than a consumer; 12.7 percent of dealers and 4.9 percent of consumers receive a prison sentence of more than five years.

### 3.2.2 Arrest Rates: UCR

The probability of arrest is constructed using county level data on number of arrests for possession and sale of illicit drugs. The FBI’s annual Uniform Crime Reporting (UCR) program provides a nationwide view of crime based on statistics submitted by law enforce-

\(^{32}\)I use the median rather than the mean so that extreme outliers do not bias the statistic. For instance, some felons convicted of drug possession receive unusually long sentences. This may be due to a ‘three strikes rule’ or some factor that is unobserved in the data (felon history is not reported in NJRP). I use the county level sentence data instead of state law mandates. Given that a dealer is considering his expected penalty, I assume dealers are more likely to know and react to the sentence history of their fellow drug dealers rather than the state legal code.

\(^{33}\)Note that the enforcement variables presented are transformed according to the methods described in Appendix B. Analyzing the means and standard deviations of these transformed variables is not straightforward.
ment agencies at the city, county and state level. Table 3 presents the average number of arrests per county for each offense conditional on observing arrests for possession and sale of cocaine. The average number of arrests for sale increases as more counties become involved in the UCR program. For example, the mean number of sale arrests increases from 60 in 1985 to 143 in 1990, and the number of counties reporting expanded by almost 50 percent during that time period. From 1990 to 2000 the number of counties reporting increases by 22 percent; however the number of arrests per county falls for both offenses. Arrests for sale decrease by 61 percent and the average number of arrests for possession decreases by 26 percent. In 2000, law enforcement agencies active in the UCR program represented 94 percent of the total population of the United States (96 percent of the population in Metropolitan Statistical areas (MSAs), 87 percent outside metropolitan areas, and 88 percent in rural counties).

The number of arrests from the UCR data acts as a numerator in the calculation of probability of arrest for each offense (possession or sale) in county $l$. 

4 Preliminary Analysis

Prior to estimating the structural model of drug dealer and consumer behavior described above, I estimate some reduced form models. For this analysis, I merged the purchase and penalty data by year and county.

Reduced form regression models are useful for understanding the crude relationships that exist in the data. The variables used are pure gram price, arrest rate and median prison or probation sentence length for possession and sale. The pure gram price of the $j^{th}$ purchase in county $l$ in year $t$ is defined as $p_{ljt}$. The enforcement variables are measured for a given county and year. I observe the arrest rate $A_{lt}$ and the median sentence length $S_{lt}$ for

---

34 The FBI does not separately report cocaine and opium related arrests. This group definition of arrests may overestimate the probability of arrest. An alternative would be to use data from another source to adjust the arrest data. The Advanced Drug Abuse Monitoring (ADAM) reports cocaine and heroin emergency room visits by geographic location. One could potentially use the ratio of cocaine to heroin emergency room visits to adjust for the arrest numbers for cocaine arrests only.

35 See Appendix B for a full discussion of how the data is used to create the probability of arrest.

36 See Appendix B for further explanation.

37 This variable is calculated by dividing the individual purchase-price by purchase-weight times purchase-purity for each observation.
possession and sale of cocaine in a given county-time combination.\textsuperscript{38} The expected penalties are constructed by multiplying the median prison sentence length (or probation term) by the arrest rate for sale or possession of cocaine.\textsuperscript{39}

Table 4 presents the signs of the coefficients from a regression of pure gram prices on the enforcement variables with year fixed effects. This regression is run separately for consumer and dealer enforcement variables (columns 3 and 4 in Table 4, respectively). The coefficient estimates on expected prison sentence for both sale and possession of cocaine are not significant. The negative coefficients on probation are significant for consumers (possession) but not dealers (sale). The fact that the results are not significantly different from zero is not surprising given that this analysis is performed with only time fixed effects and does not account for the structural factors of the model.

The coefficient estimates reported in Table 4 do not completely match the common beliefs of how prices vary with penalties. The legal penalty for possession of cocaine imposes a cost on the consumer. The conventional supply and demand model suggests a negative relationship between demand side enforcement and price (consumers leave the market in response to enforcement, resulting in a lower price). Within the search model, an increase in the consumer-side enforcement increases the cost of search and causes the reservation value \( R \) to fall.\textsuperscript{40} A decrease in the reservation value increases the probability of making a sale for any given dealer. Holding all other variables constant, when a dealer maximizes profit given the new reservation values and probability of making a sale, he chooses a higher price. Thus increased consumer enforcement puts upward pressure on pure gram price. The positive coefficient on expected prison sentence for possession in Table 4 is consistent with the search model.

The theory of risks and prices suggests that dealers need to be reimbursed for high costs or long sentences, and thus prices should be higher when dealers face long expected sentences for selling cocaine.\textsuperscript{41} Although the negative coefficient on sale-penalty may be counter-intuitive when analyzed only within the traditional supply and demand framework, the sign is consistent with dealer behavior in the search model described above. Consider

\textsuperscript{38}The median probation term is denoted \( B_{lt} \) for county \( l \) at time \( t \).
\textsuperscript{39}Calculation of the arrest rate is discussed in section 5.5 and Appendix B.
\textsuperscript{40}See Figure 3.
\textsuperscript{41}See Reuter and Kleiman (1986).
an increase in the enforcement variables for the dealer, holding everything else constant. The increased cost is incurred only if the dealer is caught selling cocaine. Maximizing profit requires increasing the probability of making a sale. Holding weight and purity constant, the dealer wants to sell his packages quickly and minimize the chances of being arrested. Thus increased dealer enforcement puts downward pressure on pure gram price, which is consistent with the negative sign on expected prison sentence for sale in Table 4.

The search model does not contradict traditional supply and demand analysis entirely. An increase in consumer side enforcement increases the search cost and causes some consumers to stop searching; the positive value they receive from consuming cocaine does not outweigh the potential cost of search. Exit of consumers results in the demand curve shifting back and decreasing price as in Figure 2. Given the positive pressure on price described above, an increase in consumer-side enforcement may cause a negative or positive change in pure gram price. If dealer-side enforcement increases the expected cost of the dealer to the point where he no longer makes a positive profit, the dealer exits the market. Dealer exit results in the supply curve shifting back and increasing price as in Figure 1. If the negative pressure on pure gram price from the increase in dealer enforcement is greater than the increase in price due to dealer exit, the equilibrium price is negatively correlated with dealer enforcement. Theory alone does not reveal the total effect of enforcement on pure gram price; it is an empirical issue.\textsuperscript{42}

Although this reduced form model is similar to the work of Bushway, Caulkins and Reuter (2003), I do not use the city-level prices prepared by Abt Associates and I use different data to represent penalties. Therefore, it is not surprising that the reduced form estimates of Table 4 differ from the results of Bushway et. al. (2003). Using an “expected cost” model the authors report negative, but not significant coefficients on dealer arrest rate and positive significant coefficients on expected sentence.\textsuperscript{43}

Table 5 presents the results when I regress the pure gram price on arrest rates. The probability of arrest plays an important role in the search model so it is important to see the

\textsuperscript{42}While these relationships can be considered intuitive, they are not straightforward to prove within the model (because $\frac{dP}{dP}$ cannot be computed analytically).

\textsuperscript{43}Bushway et. al. estimate several regression models using variations of dependent variables (retail cocaine price, retail-wholesale prices, logged price) for different time periods (1983-1996, 1990-1996). In general, the coefficients on expected prison sentence are not significantly different from zero until later time periods where the dependent variable is ‘Retail - Wholesale Prices (1990-1996)’.
general relationship that exists in the data. Consistent with the equilibrium search model, I find a negative correlation between the arrest rate for sale and pure gram price. If the probability of being arrested for sale is high, dealers want to decrease the pure gram price so they can sell the packages quickly and easily. However, the coefficient for possession is also negative, which is not consistent with the search model. If the probability of being arrested for possession is high, consumers are willing to accept less satisfactory packages to avoid being arrested on the next sample. This allows a dealer to increase the pure gram price of his package and take advantage of the low reservation values of the consumers. Given the model, I would expect a positive relationship between pure gram price and the arrest rate for possession.

These estimates also differ from the results of previous work. Using total per capita drug offense arrests as a proxy for certainty, Kuziemko and Levitt (2001) find a positive and significant coefficient for certainty on pure-gram price. I regress pure gram price on the arrest rate for possession and the arrest rate for sale separately to see if consumer-side and dealer-side enforcement have separate effects.44

Table 6 presents the coefficient estimates when purity is regressed on the expected penalties including county and time fixed effects. Recall the assumption that consumers and dealers perfectly observe purity of cocaine packages. I would expect no relationship with enforcement variables if purity is unobserved to the agents. However, there is a positive and significant relationship between purity and each of the expected sentence variables. Note that there is no direct causation implied by the mean regressions presented in the tables. A dealer may choose to produce a very pure package in order to charge a high price on few packages and not expose himself to heavy enforcement. The other possibility is that the expected penalties are increased by lawmakers in counties and years when the cocaine is very pure.45

44Kuziemko and Levitt also define retail cocaine differently. They use city-year averages obtained from cocaine purchases of five ounces or less. Five ounces translates to 140 grams of cocaine and could be considered wholesale, rather than retail level purchases. Bushway, Caulkins and Reuter restrict the retail prices to purchases made at the one ounce level or below in order to drop purchases made at the wholesale level.

45This potential endogeneity problem applies to the structural equilibrium model as well. If lawmakers change penalties in response to changes in the distribution of packages types, there would be a great deal of variation in the expected penalties from year to year. However, the null hypothesis that county penalties are stable over time is not rejected (See Appendix). I treat the county level penalties for consumers and dealers as exogenous.
Tables 7 and 8 present the correlation matrix of the expected enforcement and median enforcement variables, respectively. Note that probation is negatively correlated with prison sentence for both sale and possession. Counties with long prison sentences are likely to have short probation terms (and vice versa). Prison sentence is positively correlated across sale and possession (0.3756 and 0.8333), implying that counties are consistent in assigning sentences across drug offences. Correlation of probation terms is very low (0.1991 and 0.13152).

5 Estimation

The parameters of the model are estimated by maximum likelihood using the data described in Section 3. The parameters to be estimated include the distributional parameters of preferences $\theta \sim \Phi(\bar{\theta}, \Omega)$ and the distributional parameters of the vector of errors $e \sim F_e(\bar{e}, \Sigma)$. I also estimate the vector of coefficients on the penalty data $\Delta = (\delta)$. Conditional on parameter values, I solve for the equilibrium of the model using equations (3) and (11). The likelihood function is then evaluated for the data given the parameter values and the solution to the model. Each package in the observed data corresponds to a profit-maximizing package type choice for some dealer. The likelihood contribution of each dealer $j$ is the likelihood of observing dealer $j$ producing his particular type of package, conditional on the probability of observing dealer $j$.

The following section specifies the functional form for the distribution of consumer preferences as well as the distribution of structural errors. Next I present the equilibrium solution algorithm and describe how the probability of making a sale is calculated. I then derive the likelihood function. The section concludes with the specification of the probability of arrest and a brief discussion of identification.

5.1 Empirical Specification

Recall consumers evaluate the consumption of cocaine relative to the outside option. All consumers have preferences for cocaine, however a consumer with a negative $\beta$ does not participate in the market. I assume an i.i.d. normal distribution for consumer preferences: $\theta \sim N(\bar{\theta}, \Omega)$ where $\bar{\theta}$ and $\Omega$ are parameters to be estimated.
Each dealer has a dealer-specific vector of errors that determine the profit-maximizing choice of package; \( \varepsilon \) and \( \omega \) appear in the number of packages function (equation (6)) and \( \eta \) represents variation in penalties received by dealers. The vector of structural errors is assumed to be distributed: 

\[
\mathbf{e} = \begin{pmatrix} \varepsilon \\ \omega \\ \eta \end{pmatrix} \sim F_e(\mathbf{e}, \Sigma) \text{ with } \mathbf{e} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \text{ and } \Sigma = \\
\begin{pmatrix} \sigma_{\varepsilon} & \sigma_{\varepsilon\eta} & \sigma_{\varepsilon\omega} \\
\sigma_{\varepsilon\omega} & \sigma_{\omega} & \sigma_{\varepsilon} \\
\sigma_{\varepsilon\eta} & \sigma_{\eta} & \sigma_{\eta\omega} \end{pmatrix}.
\]

The \( \sigma \)'s are parameters to be estimated.

5.2 Model Solution

For each set of parameter values \((\hat{\theta}, \hat{\Sigma} \text{ and } \hat{\Delta})\), the model is solved for the equilibrium distribution of packages \(F_E^k(x)\) and reservation values \(F_R^k\). The algorithm takes the following form:

1. Create a grid of values for \(\hat{\theta}\) from the distribution \(\Phi(\theta)\). Each value of \(\hat{\theta}\) represents a consumer. Create a grid of values of \(\mathbf{e}^* = \begin{pmatrix} \varepsilon^* \\ \omega^* \\ \eta^* \end{pmatrix}\) from the error distribution \(F_e\).

Each \(\mathbf{e}^*\) represents a dealer.

2. Begin at \(k = 1\). Assume a distribution of package types for \(F^k(x)\).

3. Using the package types defined by \(F^k(x)\) and the parameter \(\hat{\theta}\), calculate the reservation value \(R(\hat{\theta})\) for each consumer (i.e., for each gridpoint). This gives a distribution of reservation values – \(F_R^k\).

4. Given the distribution of reservation values \(F_R^k\) from Step 3 and the data on penalties \((\gamma W)\) calculate the profit-maximizing choice of package \((p, q, s)\) for each dealer (i.e., for each gridpoint of dealer-specific errors \(\mathbf{e}^*\)). The profit-maximizing solution for each dealer (gridpoint) creates a new distribution of packages \(F^{k+1}(x)\).

5. Set \(k = k + 1\)
6. Go to Step 2 unless convergence occurs. (i.e., use $F_{k+1}(x)$ in step (2) to generate a new distribution of reservation values $F_{R_{k+1}}(x)$). I test for convergence to the solution by checking if $F_{k+1}(x) - F_k(x) < \epsilon$ where $\epsilon$ is some threshold close to zero.

5.3 Probability of Making a Sale, $P^j$

When the model is solved for equilibrium using the algorithm described above, the dealer chooses to produce the profit-maximizing type of package. For each possible package, a dealer computes the probability of selling that package. Recall equation (7) that defines the probability $P_j$ that dealer $j$ sells his package $x_j$:

$$P_j = \int_{\{\theta: V(x_j, \theta) > R(\theta)\}} d\Phi(\theta)$$

The probability cannot be computed analytically so I compute $P^j$ numerically. I allow each element of $\theta$ to take on five values. For each package $x_j$, I evaluate the function $V(x_j, \theta) - R(\theta)$. Iterating over values of $\theta$ defines the set of consumers that purchase the given package. See the Appendix for a more complete description of the algorithm for computing $P^j$.

5.4 The Likelihood Function

Solving the model results in an equilibrium distribution of reservation wages and package types. Recall the model has been solved conditional on a guess of the parameters including $\hat{\Sigma}$ the covariance matrix which defines the distribution of errors $F_e$. Using the equilibrium distributions $F_{R_E}$ and $F_E(x)$, I calculate the probability of selling each package type $P^j$. Given the probability of sale for each package, I find the draws of the error vector that would be necessary to observe the packages in the data as profit-maximizing choices (i.e., what does each dealer need to realize in order to make the choice that is observed in the data). Each dealer $j$ has an error draw $\hat{\epsilon}_j$ such that package $j$ is his profit-maximizing choice. Finding $\hat{\epsilon}_j$ such that the first order conditions for dealer $j$ are zero is equivalent to minimizing the square of the vector of first order conditions with respect to $\epsilon_j$.

The likelihood contribution of each dealer is the likelihood that dealer $j$ is observed choosing to produce package $x_j$. However, dealers vary in the number of approaches $D_j$ they make in order to sell all $N_j$ packages. A dealer with a high number of approaches is more
likely to be seen in the data, whereas a dealer with a small number of approaches is less likely to be observed because he sells his packages and leaves the market. Thus the sample under-represents dealers with a small number of packages. Recall the number of approaches $D_j$ is a random variable with expected value $E(D_j) = \frac{N_j}{P_j}$ conditional on the dealer specific error vector $\bar{e}_j$ that determines $N_j$ and $P^j$. The density of the error resulting in package $j$ being the profit-maximizing choice is $f_{e}(\bar{e}_j)$.

Thus, the likelihood function for all dealers can be expressed as:

$$L = \prod_j \left[ \frac{f_{e}(\bar{e}_j)E(D_j|\bar{e}_j)}{\int f_{e}(e)E(D_j|e)de} \right]$$

(12)

where $\int f_{e}(e)E(D_j|e)de$ is the proportion of dealers of type $j$ in the market.

5.5 Probability of Arrest

It is necessary to compute a reasonable measure of the probability of arrest for possession and sale to estimate the structural model described in Section 2.46

The probability of arrest for possession of cocaine in county $l$ is defined to be $Pr(A_P^l)$ and the probability of arrest for sale of cocaine in county $l$ is defined to be $Pr(A_S^l)$.

$$Pr(A_P^l) = \frac{\text{Number of Arrests for Possession in county } l \text{ for year } t}{\text{Total Number of Consumer searches conducted in county } l \text{ for year } t}$$

(13)

$$Pr(A_S^l) = \frac{\text{Number of Arrests for Sale in county } l \text{ for year } t}{\text{Total Number of Dealer approaches in county } l \text{ for year } t}$$

(14)

Each time a consumer samples a package she is met by a dealer approach. The total number of approaches by dealers is equal to the total number of consumer samples. I define the denominator in both equation (13) and equation (14) to be the number of encounters $E$ between consumers and dealers.

There are two ways to calculate the number of encounters $E$: use either the dealer side of the model or the consumer side of the model. Both methods require a measure of the market size in each county $l$. However, there does not exist any exogenous measures of the number of dealers across counties. There is limited information on cocaine consumers from

46DeSimone (2001) proxies the probability of arrest using the ratio of arrests to reported offenses given in the UCR data. DeSimone and Farrelly (2001) proxy the probability of arrest with a variable equaling the number of arrests in the state divided by the number of drug users in the state that year. They construct the denominator using responses to the National Household Survey on Drug Abuse (NHSDA) and Census data on state population.
the National Household Survey on Drug Abuse (NHSDA).\footnote{The National Household Survey on Drug Abuse (NHSDA) is a yearly survey conducted by the Office of Applied Studies at the Substance Abuse and Mental Health Services Administration. The NHSDA is designed to produce drug and alcohol use incidence and prevalence estimates and report the consequences and patterns of use and abuse in the general U.S. civilian population aged 12 and older. Questions include age at first use, as well as lifetime, annual, and past-month usage for a multitude of drugs, including cocaine and crack. Demographic data include gender, race, age, ethnicity, educational level, and population density.} I use the consumer-side method to construct $E$ for use in both the probability of arrest for possession $\Pr(A_P^l)$ and probability of arrest for sale $\Pr(A_S^l)$.

Recall that the probability a consumer is successful in her search for cocaine is defined by the distribution of her values given the distribution of package types. The probability that a consumer fails to purchase the offered package $x$ is $G_d(R(\theta)) = \Pr(V(x, \theta) < R(\theta))$. The number of searches a consumer makes before finding a package above her reservation value is given by $\frac{1}{G_d(R(\theta))}$. This reveals the number of searches for a given consumer with preferences $\theta$ but not the total number of searches for all consumers in market $l$. The average number of searches per consumer is found by integrating over the distribution of preferences $\theta$: $\int \frac{1}{G_d(R(\theta))} d\Phi(\theta)$.\footnote{Recall that a consumer with a negative $\beta$ is not considered to be in the market. A consumer with a positive value of $\beta$ who chooses not to purchase (because the search cost outweighs the maximum possible value) is defined as not searching.}

I define $M^C_l$ to be a measure of the size of the consumer market in county $l$.\footnote{Appendix B describes how this measure of market size $M^C_l$ is created.} The number of encounters $E$, is calculated as the average number of searches per consumer times the total number of consumers:

$$E = M^C_l \frac{1}{G_d(R(\theta))} \Phi(\theta)$$

Given the number of encounters $E$, the probability of arrest for possession $\Pr(A_P^l)$ and sale $\Pr(A_S^l)$ are computed using equations (13) and (14).

## 5.6 Identification

I make several identifying assumptions in order to estimate the parameters of the model. On the consumer side of the model, I assume the distribution of consumer preferences is a national distribution. The covariation of observed search costs (possession penalties) and arrest rates for possession across counties with multiple package types within a county
identifies the distributional parameters of preferences \( \theta \sim \Phi(\theta, \Omega) \).\(^{50}\)

On the dealer side of the model, I assume all dealers receive the same average wholesale stock of cocaine \( K \) and only produce one type of package.\(^{51}\) The assumption of identical wholesale stock for all dealers is necessary to identify the number of packages \( N_j \), given only data on purity \( s_j \) and weight of a single package \( q_j \). The variation in firm costs (sale penalties), arrest rates for sale and the multiple observations of packages within a county identify the parameters of the error distribution \( F_e \).

I also use the assumptions of unit demand by consumers and static profit-maximization by dealers to identify the model. There is no data on the quantities sold by each dealer. Given that the choices of price, weight and purity are the profit-maximizing choices at equilibrium, I calculate the probability of selling each potential package. Consumers demand only one package so the implied demand (probability of making a sale) from the consumer side of the model is a function of the choices of the dealers. I substitute the probability of making a sale as a function of the choice variables into the dealer profit function and identify both sides of the market using the first order conditions of the dealer.

6 Results

7 Policy Questions

The principal advantage of estimating a structural model is that it allows the econometrician to simulate changes in the underlying economic environment. Given the packages that exist in a county and the expected penalty for consuming and selling cocaine, I estimate the parameters of the distribution of consumer preferences \( \Phi(\theta) \) and the parameters for the distribution of the errors \( F_e \). These distributions characterize the market for cocaine and are necessary components to understand the impact of potential changes in policy. Following estimation of the structural parameters, I simulate a change in the penalty structure. Policymakers often discuss how the market for cocaine would respond if the current sentences for sale of cocaine were increased or decreased. Because the model is based on the underlying

\(^{50}\)Recall that most equilibrium search models require duration data to identify the model. Because I observe search costs as data and use restrictions implied by the equilibrium, duration data is not necessary.

\(^{51}\)Recall that I assume dealers produce multiple packages of identical weight, purity level, and price to sell in the retail market.
optimization of both consumers and dealers, I can simulate changes in the length of sentences for both possession and sale of cocaine.

If policymakers increase the current sentences for sale of cocaine, dealers adjust the types of profit-maximizing packages they choose to produce. Consumers react to the new distribution of available packages in the market by changing reservation values and purchasing decisions. If the penalty for possession of cocaine is increased, the reservation value of every consumer will decrease, dealers will adjust production along the three dimensions of package type and the distribution of packages will change. Given the change in package type distribution, I can predict how total cocaine consumption changes in response to the alteration of the market. Using the structural estimates, I can discuss which type of enforcement (consumer-based or dealer-based) has the largest impact on the market for cocaine.

Another policy experiment I intend to conduct increases the probability of arrest for sale of cocaine. Increasing the frequency of enforcement is analogous to increasing the number of police patrolling the streets. Again, this policy change affects the profit-maximizing decisions of cocaine dealers and a new consumer-dealer equilibrium is reached. This new equilibrium may have different levels of consumption and result in different equilibrium package type distributions.

8 Conclusion

This paper develops and estimates an equilibrium search model in which both drug dealers and consumers consider the legal penalties of selling and holding illicit drugs when making their respective optimization decisions. Specifically, consumers choose a reservation value and purchasing strategy that incorporates the median prison sentence for possession of cocaine. Drug dealers maximize a profit function that incorporates the median prison sentence for sale in their respective counties. The model allows dealers to choose the purity level of cocaine they sell as well as the weight and price of their package.

Debate over prohibition of narcotics is hardly a new issue. Use and production of cocaine and other illicit substances have been major policy concerns for the last two decades. Enforcement of current drug laws consumes a large share of the federal budget and drug offenders make up the largest share of inmates in federal prisons. While there is a growing body
of research on the relationship between enforcement and retail price of cocaine, preliminary
data analysis underscores the need for a structural equilibrium model of cocaine production
and consumption. This paper adds a necessary equilibrium component to research on the
drug market. Consumers and dealers are modelled as making interactive decisions, generat-
ing structural parameter estimates. I use these estimates to simulate policy changes allowing
for equilibrium effects.
References


[27] System to Retrieve Information from Drug Evidence (STRIDE), Drug Enforcement Agency (DEA). Made available through the Freedom of Information Act. Request Number: 01-1799-F


Appendix

A.1 Proof that Reservation Value, $R(\theta)$ exists

Package characteristics $(p, q, s)$ are distributed $F$ and $V(x, \theta)$ is a function mapping draws of $(p, q, s)$ into a scalar value for each person. There is a lower bound to $(p, q, s)$ (i.e., purity, price and weight cannot be negative) which implies some lower bound $V(\theta)$ to the distribution of values for each person. There also exists an upper bound $\bar{V}(\theta)$ that varies over consumers.

Recall $H(V')$ is defined as the expected benefit to search given a current package with value $V'$. In order to prove the existence of reservation value $R(\theta)$ for a given consumer, the following must hold:

- $H(V')$ is continuous over the support of $G(V)$.
- $\lim_{V' \to \bar{V}} H(V') = 0$ (As the value of current package approaches the upper bound of possible values, the expected benefit from searching again approaches zero).
- $\lim_{V' \to V} H(V') > c$ (If the value of the current package is the lower bound of all values, the expected benefit from searching again is greater than the cost of searching for the next package.\textsuperscript{52}
- $H(V')$ is decreasing in $V'$.

The fourth point can be shown to be true by taking the derivative of $H(V')$ with respect to $V'$. Recall $H(V') = \int_{V'}^{V}(V - V')dG(V) = \int_{V'}^{V} VdG(V) - V'G(V) + V'G'(V')$.

\[
\frac{dH(V')}{dV'} = \frac{d}{dV'} \left\{ \int_{V'}^{V} VdG(V) \right\} - G(V) + V' \frac{dG(V)}{dV'} + G(V') + V'G'(V')
\]

\[
= \frac{d}{dV'} \left\{ \int_{V'}^{V} VdG(V) \right\} - G(V) + G(V') + V'G'(V')
\]

Following Leibnitz' Rule:

\[
\frac{d}{dV'} \left\{ \int_{V'}^{V} VdG(V) \right\} = \int_{V'}^{V} V'G'(V)dV = -V'G'(V')
\]

\textsuperscript{52}Note: If the expected benefit from the next package is less than the cost of search (i.e., $H(V) < c$), then the consumer will not be in the market. If $H(V) = c$, the consumer is indifferent between being in or out of the market.
\[
\frac{dH(V')}{dV'} = -V'G'(V') - G(V) + G(V') + V'G'(V') = -G(V) + G(V')
\]

Recall that \( G() \) is a cdf and is increasing in its argument. Also \( \bar{V} \) is the upper bound of \( V() \) which implies \( V' \leq \bar{V} \). It follows that

\[
\frac{dH(V')}{dV'} \leq 0.
\]

Therefore, the function \( H(V') \) is decreasing in \( V' \).

Given that these conditions hold for each consumer in the market, draw \( c \) and \( H(V') \) in the same curve (see Figure 3). The reservation value \( R \) occurs at the intersection.\(^{53}\)

### A.2 Deriving the Profit Function

Recall in section 3.6 that the profit function is expressed as

\[
E(\Pi_j) = \sum_{D_j} (1 - Pr_{AS}^{D_j}) [Pr(D_j)] \ast p_j \ast N_j - \left[ 1 - \sum_{D_j} (1 - Pr_{AS}^{D_j}) [Pr(D_j)] \right] \ast [\gamma W + \eta_j] - T
\]

and that \( \sum_{D_j} (1 - Pr_{AS}^{D_j}) [Pr(D_j)] \) can be expressed as \( E(1 - Pr_{AS}^{D_j}) \). It follows that

\[
\sum_{D_j} (1 - Pr_{AS}^{D_j}) [Pr(D_j)] = E(1 - Pr_{AS}^{D_j}) = E \{ \log(1 - Pr_{AS}^{D_j}) \}
\]

Using the moment generating function for the negative binomial distribution, the term on the left hand side of 16 can be expressed as

\[
MD_j \{ \log(1 - Pr_{AS}^{D_j}) \}
\]

\[
= \frac{(P_j^{\log(1 - Pr_{AS})})^N_j}{[1 - (1 - P_j)^{\log(1 - Pr_{AS})}]^N_j} \cdot \frac{[P_j(1 - Pr_{AS})]^N_j}{[1 - (1 - P_j)(1 - Pr_{AS})]^N_j}.
\]

Replace the left hand term in equation (16) with equation (17) and rewrite the expected profit function:

\[
E(\Pi_j) = \frac{[P_j(1 - Pr_{AS})]^N_j}{[1 - (1 - P_j)(1 - Pr_{AS})]^N_j} \ast p_j \ast N_j - [1 - \frac{[P_j(1 - Pr_{AS})]^N_j}{[1 - (1 - P_j)(1 - Pr_{AS})]^N_j}] \ast [\gamma W + \eta_j] - T
\]

\(^{53}\)It may be possible that there is no intersection. This is the case where \( H(\bar{V}) < c \), and the consumer is not in the market.
A.3 Proof that $\Pr(V > R)$ is monotone in $\theta$

For every package $x_j^*$, there exists a $\theta^*$ representing the preferences of the marginal consumer who chooses to purchase $x_j^*$. The set of consumers is divided into the set that purchase package $x_j^*$ and the set that do not purchase $x_j^*$. This is equivalent to proving that $\Pr(V > R)$ is monotone in $\theta$. I begin by introducing new notation that is relevant only to the proof at hand.

Let $Y = f(V, \theta)$ where $Y \sim H(y)$. It follows that $V = f^{-1}(Y, \theta)$. The distribution of $V$ is:

$$
G(v) = \Pr(V \leq v) = \Pr(f^{-1}(Y, \theta) \leq v) = \Pr(Y \leq f(v, \theta)) = H(f(v, \theta))
$$

$$
g(v) = G'(v) = h(f(v, \theta))f'(v, \theta)
$$

(19)

To find $\frac{d}{d\theta}[\Pr(V > R)]$:

$$
\frac{d}{d\theta}[\Pr(V > R)] = \frac{d}{d\theta}[1 - H(f(R, \theta))] = -h(f(R, \theta)) \frac{df(v, \theta)}{d\theta} |_{v=R}
$$

$$
= -h(f(R, \theta)) \left[ \frac{\partial f(v, \theta)}{\partial v} |_{v=R} \frac{\partial R}{\partial \theta} + \frac{\partial f(v, \theta)}{\partial \theta} |_{v=R} \right]
$$

(20)

From the Implicit Function Theorem:

$$
\frac{\partial f(v, \theta)}{\partial \theta} |_{v=R} = -\frac{\partial f(v, \theta)}{\partial v} |_{v=R} \frac{\partial R}{\partial \theta} + \frac{\partial f(v, \theta)}{\partial \theta} |_{v=R}
$$

(21)

Recall that $\frac{\partial f^{-1}(y, \theta)}{\partial \theta}$. Equation (20) can be written as:

$$
\frac{d}{d\theta}[\Pr(V > R)] = -h(f(R, \theta)) \frac{\partial f(v, \theta)}{\partial v} |_{v=R} \left[ \frac{\partial R}{\partial \theta} - \frac{\partial f^{-1}(y, \theta)}{\partial \theta} |_{v=R} \right]
$$

(22)

A.3.1 Find $\frac{dR}{d\theta}$

To find $\frac{dR}{d\theta}$, begin by evaluating the conditions necessary for the existence of a reservation value. Let $Q$ be defined as the value of holding a current package. The optimal stopping rule was outlined in an earlier section. The value of using this strategy is:

$$
Q = (1 - P) \int_R^\infty Vh(f(v, \theta)f'(v, \theta)dv + PJ + H(f(R, \theta))Q
$$

$$
= \frac{(1 - P)}{1 - H(f(R, \theta))} \int_R^\infty Vh(f(v, \theta)f'(v, \theta)dv + \frac{PJ}{1 - H(f(R, \theta))}
$$
where $P$ is the probability of being arrested for possession and $J$ is the jail time/penalty for possession.$^{54}$

To find the optimal reservation value, maximize with respect to $R$:

$$
\frac{dQ}{dR} = \frac{(1-P)}{[1-H(f(R,\theta))]^2} \cdot h(f(R,\theta))f'(R,\theta) \int_R^\infty V h(f(v,\theta))f'(v,\theta) dv + \frac{P0J}{[1-H(f(R,\theta))]^2} \cdot h(f(R,\theta))f'(R,\theta) + \frac{(1-P)}{[1-H(f(R,\theta))]} [-Rh(f(R,\theta)f'(R,\theta)]
$$

$$
= \frac{h(f(R,\theta))f'(R,\theta)}{[1-H(f(R,\theta))]^2} \left[ (1-P) \int_R^\infty V h(f(v,\theta))f'(v,\theta) dv - (1-P)R[1-H(f(R,\theta))] + PJ \right]
$$

Thus, the necessary condition for $R$ is:

$$
Z(R,\theta) = (1-P) \int_R^\infty (V-R)h(f(v,\theta))f'(v,\theta) dv + PJ = 0 \quad (23)
$$

Integrate $Z(R,\theta)$ using integration by parts. Note: The first term in the brackets goes to zero.

$$
Z(R,\theta) = (1-P) \left[ -(V-R)[1-H(f(v,\theta))] R \int_R^\infty [1-H(f(v,\theta))] dv \right] + PJ
$$

$$
= (1-P) \int_R^\infty [1-H(f(v,\theta))] dv + PJ = 0 \quad (24)
$$

Using the implicit function theorem, $\frac{\partial R}{\partial \theta} = -\frac{dR}{d\theta}$. First:

$$
\frac{dZ}{d\theta} = (1-P) \int_R^\infty -h(f(v,\theta)) \frac{df(v,\theta)}{d\theta} dv \quad (25)
$$

Again with use of the implicit function theorem, $\frac{df(v,\theta)}{d\theta} = -\frac{df(v,\theta)}{dv} \cdot \frac{dv}{d\theta}$ so (25) is rewritten:

$$
\frac{dZ}{d\theta} = (1-P) \int_R^\infty h(f(v,\theta)) \frac{df(v,\theta)}{dv} dv \quad (26)
$$

Next:

$$
\frac{dZ}{dR} = (1-P) \{-[1-H(f(R,\theta))] \} \quad (27)
$$

Given this, $\frac{\partial R}{\partial \theta}$ can be written as

$$
\frac{\partial R}{\partial \theta} = -\frac{dZ}{d\theta} = \frac{(1-P) \int_R^\infty h(f(v,\theta)) \frac{df(v,\theta)}{dv} dv}{(1-P)[1-H(f(R,\theta))]} \quad (28)
$$

$^{54}$Note: The notation $Q, P$, and $J$ differ from the notation used in the model described earlier. This notation refers only to the general proof at hand.
Recall \( g(v) = h(f(v, \theta))f'(v, \theta) \). Equation (28) is written:
\[
\frac{dR}{d\theta} = E \left[ \frac{dV}{d\theta} | V > R \right]
\]  \hspace{1cm} (29)

### A.3.2 Plug In \( \frac{dR}{d\theta} \)

Now that we have \( \frac{\partial R}{\partial \theta} \), we can sign (22).
\[
\frac{d}{d\theta} \left[ \Pr(V > R) \right] = -h(f(R, \theta)) \left. \frac{\partial f(v, \theta)}{\partial v} \right|_{v=R} \times \left[ \left. \frac{\partial R}{\partial \theta} - \left. \frac{\partial f^{-1}(y, \theta)}{\partial \theta} \right|_{v=R} \right] \\
= -h(f(R, \theta)) \left. \frac{\partial f(v, \theta)}{\partial v} \right|_{v=R} \times \left[ \left( \frac{dv}{d\theta} | V > R \right) - \left. \frac{\partial v}{\partial \theta} \right|_{v=R} \right] \\
= (-) \times (?) \times (+)  \hspace{1cm} (30)
\]

Meaning that the sign of \( \left. \frac{\partial f(v, \theta)}{\partial v} \right|_{v=R} \) will determine the sign of \( \frac{d}{d\theta} \left[ \Pr(V > R) \right] \). As long as \( \left. \frac{\partial f(v, \theta)}{\partial v} \right|_{v=R} \) is non-zero (and always positive or always negative), we have proven that \( \Pr(V > R) \) is monotone in \( \theta \).

### A.4 Equilibrium

The following lists conditions necessary for the existence of an equilibrium. Once these conditions are shown to hold, equilibrium follows from Schauder’s fixed point theorem.

Given:

- Define \( \Upsilon \) as the space of distributions of reservation values, \( F_R \).

- \( \Psi_R(\Psi_x(F_R, F_e), \Phi) \) is the function that maps \( F_R \) into a subset of the space \( \Upsilon \), which I define as \( \Upsilon^* \).

Necessary Conditions:

- \( \Upsilon \) is a non-empty, convex set in a Banach space.

- The set that \( F_R \) is mapped into from \( \Upsilon \) by the function \( \Psi_R(\Psi_x(F_R, F_e), \Phi) \) must be a compact subset of the space \( \Upsilon \).

  - \( \Upsilon^* \subset \Upsilon \).

  - \( \Upsilon^* \) is convex and compact.

- \( \Psi_R(\Psi_x(F_R, F_e), \Phi) \) is a continuous mapping of \( F_R \).
A.5 Calculation of Probability $P^j$

The probability of making a sale for each dealer ($P^j$ for a given package) is defined as:

$$P^j = \int \frac{d\Phi(\theta)}{\{\theta: V(x, \theta) \geq R(\theta)\}}$$

which cannot be computed analytically. The following algorithm is used to calculate the probability of making a sale conditional on a given package type.

Recall $\theta = (\beta, \lambda_q, \lambda_s, \alpha)$ and $V_{ij} = \beta_i q_j \lambda_q \lambda_s - \alpha_i p_j$.

1. Recall $\theta_m \sim N(\overline{\theta}_m, \Omega_m)$. (Assume some values for $\overline{\theta}_m, \Omega_m$). Pick the values of $\theta^i_m$ where $i = 1, 5$ such that the integral under the normal density function is the same for each cell (i.e., the 5 points form quartiles). Begin with $\theta_m = \alpha$. Let $\theta = \Gamma y + \overline{\theta}$ where $y \sim N(0, 1)$ and $\Gamma$ is the Cholesky decomposition matrix of $\Omega$. Using standard normal tables, the possible values for the elements of $y$ are: $-4, -0.675, 0, 0.675, 4$.

2. Condition on $y^1_m$, transform the $/m$ elements by $\Gamma y + \overline{\theta}$. Ideally, I’d like to find $\theta^i_m$ such that $V(\theta, x) = R(\theta)$. However, the reverse transformation only holds with certainty if $V(\theta, x) - R(\theta)$ is monotonic in $y$. Since monotonocity cannot be proven analytically, I use the following strategy to find $y^1_1$:

   (a) i. Begin with $y^1_1$ (all values of other $y$ elements at initial level), transform by $\Gamma y + \overline{\theta}$, and evaluate $H(\theta) = V(\theta, x) - R(\theta)$.

   ii. Repeat for each node of $y_1$.

   iii. If the sign changes between node, assume there is only one crossing point.

   iv. Find crossing point by interpolation. That point will be $y^1_1$. Note: There may be multiple sign changes along $y_1$ nodes. For now, limit possibilities to 1. Return to this issue later.

   v. If there are no sign changes, go on to the next value of $y^2_1$.

3. Keep $y_4$, and $y_3$ at their initial values, but now repeat step(2) for $y^i_1 i = 2, 5$.

   (a) Connect the points between each $y^i_1$. Note that the superscript $i$ on $y^i_1$ represents which $i$-value of $y_2$ we are at. The equation of the line connecting the two $y^i_1$
points will be:

\[ f_2(y_2) = y_1 = y_1^* + \left( \frac{y_{1(i+1)^*} - y_{1(i)^*}}{y_{2(i+1)} - y_{2(i)}} \right) y_2 \quad \forall \ i = 1, 4 \]  

(31)

(b) If in 2 dimensions, the next step would be to find the area of the part of the box above the line connecting the two \( y_1^* \) for \( i = 0, 4 \). Integrate under the lines for each cell down to the lower limit of \(-4\). \( P^2 \) represents the probability of making a sale when there are only 2 elements of \( \theta \):

\[
P^2 = \frac{1}{64} \sum_{i=1}^{4} \left[ \int_{y_{2(i)}}^{y_{2(i+1)}} \left[ f(y_2) - (-4) \right] dy_2 \right]
\]

\[
= \frac{1}{64} \sum_{i=1}^{4} \frac{y_{2(i+1)} - y_{2(i)}}{2} \left( y_1^* + y_{1(i+1)^*} + 8 \right)
\]

Take the ratio of the shaded area of the boxes to the area of the full square \((8 \times 8 = 64)\). That fraction represents the probability of being in that box.

4. Now, keep \( y_4 = y_4^1 = -4 \), and extend to the three dimensional case. Begin by setting \( y_3 = y_3^2 = -0.675 \) and repeat steps 2-3(b). Instead of having 2-dimensional boxes there are cubes and some cubes are sliced.

(a) Connect the line defined in \( y_3 = y_3^1 \)-space to the line defined in the \( y_3 = y_3^i \) \((i = 2, 5)\)-space. Rewrite (31) as: \( \tilde{y}_1 = a_{il} + b_{il} \cdot y_2; \{ y_2 \in (y_{2(i)}, y_{2(i+1)}) \}, \) for \( i = 1, 4 \} \). The \( l \) subscripts on \( a, b \) represent the segment of \( y_3 \) and the \( i \)-subscripts on \( a, b \) represent which subset of \( y_2 \) the line is defined over. \( i.e., \) the initial equation in the first cube would be written as \( \tilde{y}_1 = a_{11} + b_{11} \cdot y_2, \) where \( y_2 \in (y_{21}, y_{22}) \) and \( y_3 \in (y_{31}, y_{32}) \).

\[ a_{il} = y_{1^*} = \text{number (found in step 2 for } y_{2i}^*) \]

\[ b_{il} = \left( \frac{y_{1(i+1)^*} - y_{1(i)^*}}{y_{2(i+1)} - y_{2(i)}} \right) = \text{number} \]

\[ \tilde{y}_1^i = a_{il} + b_{il} \cdot y_2 \]

Connect the points on \( \tilde{y}_1^i = a_{il} + b_{il} \cdot y_2 \) with the points on \( y_1 = a_{1(i+1)} + b_{1(i+1)} \cdot y_2 \). (Along \( y_3 \)-axis).
5. Recall that steps (1)-(4) have all been done for the volumes in steps (1)-(4). Then integrate out over values of \( j = 2 \).

(b) For a given value of \( y'_2 \), define the equation of the line connecting \( \tilde{y}_i^l \) with \( \tilde{y}_i^{(l+1)} \).

\[
\text{Slope} = \frac{\tilde{y}_i^{(l+1)} - \tilde{y}_i^l}{\tilde{y}_3^{(l+1)} - \tilde{y}_3^l} = d_{il}
\]

\[
\text{Intercept} = \tilde{y}_i^l = c_{il}
\]

\[
\text{LineEq} : \tilde{y}_i^l = \tilde{y}_i^l + \left( \frac{\tilde{y}_i^{(l+1)} - \tilde{y}_i^l}{\tilde{y}_3^{(l+1)} - \tilde{y}_3^l} \right) y_3
\]

\[
f_3(y_3) = \tilde{y}_i^l = c_{il} + d_{il} y_3
\]  

(Note : Conditional on a value of \( y'_2 \))

which means that the area of the trapezoid defined by the space \( y'_3 \) to \( y'^{l+1}_3 \) can be written as:

\[
\text{Area(trapezoid} - l) = \int_{y'_3}^{y'^{l+1}_3} [c_{il} + d_{il} y_3] dy_3.
\]

This trapezoid exists for every value of \( y_2l \) (and the value of \( \tilde{y}_il \) depends on the value of \( y_2l \) from equation (31). The volume of the trapezoidal shape in the \( i^{th} \) subset of \( y_2 \) and the \( l^{th} \) subset of \( y_3 \) can be written as:

\[
\text{Vol} = \int_{y'_2}^{y'^{l+1}_2} \int_{y'_3}^{y'^{l+1}_3} \left[ \tilde{y}_i^l - (-4) \right] dy_3 dy_2
\]

where \( \tilde{y}_i^l = a_{il} + b_{il} \ast y_2 \).

(c) \( P^3 \) represents the probability of making a sale when \( \theta \) has three elements.

\[
P^3 = 4 \sum_{l=1}^{4} \sum_{i=1}^{4} \int_{y'_2}^{y'^{l+1}_2} \int_{y'_3}^{y'^{l+1}_3} \left[ \tilde{y}_i^l + \left( \frac{\tilde{y}_i^{(l+1)} - \tilde{y}_i^l}{\tilde{y}_3^{(l+1)} - \tilde{y}_3^l} \right) y_3 + 4 \right] dy_3 dy_2
\]

\[
= \frac{1}{512} \sum_{l=1}^{4} \sum_{i=1}^{4} \left( \frac{y'^{l+1}_3 - y'_3}{2} \ast \frac{y'^{l+1}_2 - y'_2}{2} \right) \left( y'^{l\ast}_1 + y'^{(l+1)\ast}_1 + y'^{(l+1)\ast}_1 + y'^{(l+1)(l+1)\ast}_1 + 16 \right)
\]

5. Recall that steps (1)-(4) have all been done for \( y_4 = y'_4 = -4 \). Now repeat for \( y_4 = y'^{j}_4 \), \( j = 2, 5 \). Hold \( y_2 \) constant and repeat the same procedure as what was used to create the volumes in steps (1)-(4). Then integrate out over values of \( y_2 \).

(a) Hold \( y_2 \) constant at \( y'^1_2 \). The equation for the line relating \( y_1 \) to \( y_3 \) is \( y_1 = c_{i1} + d_{i1} y_3 \). I add some notation to show which value of \( y_4 \) is being represented. The equation of the lines for \( y_4 = y'^j_4 = y'^1_4 \) is: \( \tilde{y}_i^{j} = c_{ilj} + d_{ilj} y_3 \), \( \tilde{y}_i^{11} = c_{i11} + d_{i11} y_3 \). This line is
then connected with the line that has been created for \( y_i' = y_i'^2 : y^1_{i12} = c_{112} + d_{112}y_3 \).

Recall:

\[
d_{i,j} = \left( \frac{-y_i^{(l+1)} - y_i^{(l+1)}}{y_i^{(l+1)} - y_i^l} \right)
\]

\[
c_{i,j} = \tilde{y}_{i}^{lij}
\]

I connect the lines defined on the \( y_i^l \) plane and the line on the \( y_i^2 \) plane:

\[
\text{Slope } = f_{i,j} = \left( \frac{-y_i^{(l+1)} - y_i^{(l+1)}}{y_i^{(l+1)} - y_i^l} \right)
\]

\[
\text{Intercept } = c_{i,j} = \tilde{y}_{i}^{lij}
\]

\[
\text{LineEq } : f_4(y_{4}) = y_{i}^{lij} = e_{i,j} + f_{i,j}y_4.
\]

The area for the trapezoid created by this new line connecting lines across the \( y_4 \)-space would be:

\[
\text{Area}(\text{trapezoid } - j) = \int_{y_4}^{y_4^{j+1}} [e_{i,j} + f_{i,j}y_4] \, dy_4.
\]

Integrate out over all the other dimensions. The probability in 4 dimensions is

\[
P^4 = \frac{1}{4096} \sum_{i=1}^{4} \sum_{l=1}^{4} \sum_{i=1}^{4} \left( \int_{y_2}^{y_{i+1}^{(l+1)}} \int_{y_3}^{y_{i+1}^{(l+1)}} \int_{y_4}^{y_{i+1}^{(l+1)}} [f_4(y_4) - (-4)] \, dy_4 \, dy_3 \, dy_2 \right)
\]

\[
= \frac{1}{4096} \sum_{i=1}^{4} \sum_{l=1}^{4} \sum_{i=1}^{4} \left( \int_{y_2}^{y_{i+1}^{(l+1)}} \int_{y_3}^{y_{i+1}^{(l+1)}} \int_{y_4}^{y_{i+1}^{(l+1)}} [e_{i,j} + f_{i,j}y_4 + 4] \, dy_4 \, dy_3 \, dy_2 \right)
\]

Recall that \( e_{i,j}, f_{i,j} \) are functions of \( \tilde{y}_{i}^{lij} \). The line \( \tilde{y}_{i}^{lij} \) is a function of \( c_{i,j}, d_{i,j} \) and \( y_3 \). Then \( c_{i,j}, d_{i,j} \) are functions of \( \tilde{y}_{i}^{lij} \), which is a function of \( a_{i,j}, b_{i,j} \) and \( y_2 \). The \( a_{i,j}, b_{i,j} \) are functions of \( y_i^{(i+l)} \). Once everything is plugged in, this integral can be computed analytically using iterative integration.

6. Let’s call the integral \( \bar{Y} \). Recall that the limits of integration are scalar values (the nodes chosen for \( y \) in Step (1.a)).

\[
\bar{Y} = \int_{y_2}^{y_{i+1}^{(l+1)}} \int_{y_3}^{y_{i+1}^{(l+1)}} \int_{y_4}^{y_{i+1}^{(l+1)}} \left[ e_{i,j} + f_{i,j}y_4 + 4 \right] \, dy_4 \, dy_3 \, dy_2
\]

For ease of notation, let \( y_4' = (y_4^{i+1} - y_4^l) \). Plug in the values for \( e_{i,j} \) and \( f_{i,j} \).

\[
\bar{Y} = \int_{y_2}^{y_{i+1}^{(l+1)}} \int_{y_3}^{y_{i+1}^{(l+1)}} \left[ e_{i,j}(y_4') + \frac{f_{i,j}(y_4')^2}{2} + 4 \right] \, dy_3 \, dy_2
\]

\[
= \int_{y_2}^{y_{i+1}^{(l+1)}} \int_{y_3}^{y_{i+1}^{(l+1)}} \left[ \tilde{y}_{i}^{lij}(y_4') + \left( \frac{-y_i^{(l+1)} - y_i^{(l+1)}}{y_i^{(l+1)} - y_i^l} \right) \frac{(y_4')^2}{2} + 4 \right] \, dy_3 \, dy_2
\]

40
Perform the innermost integration again with the values of $\tilde{y}_1^{ij}$. 
\[
\begin{align*}
\int_{y_2}^{y_2' + 1} \int_{y_2}^{y_2' + 1} \left[ y_1^{ilj} (y_4) + \left( \frac{y_1^{il(j+1)} - \tilde{y}_1^{ilj}}{y_4^{ij+1} - y_4^j} \right) \ast \frac{(y_4')^2}{2} + 4 \right] dy_3 dy_2 \\
= \frac{(y_4')}{2} \int_{y_2'}^{y_2' + 1} \left[ \left( c_{ilj} + c_{il(j+1)} \right) (y_3^{l+1} - y_3^l) + \left( d_{ilj} + d_{il(j+1)} \right) \frac{(y_3^{l+1} - y_3^l)^2}{2} + 8 \right] dy_2 
\end{align*}
\]
Define $y_3' = y_3^{l+1} - y_3^l$. Integrate again with the values of $c_{ilj}$ and $d_{ilj}$: 
\[
\begin{align*}
\int_{y_2}^{y_2' + 1} \left[ (c_{ilj} + c_{il(j+1)}) (y_3') + \left( d_{ilj} + d_{il(j+1)} \right) \frac{(y_3')^2}{2} + 8 \right] dy_2 \\
= \frac{(y_4')}{2} \left( y_3' \right) \int_{y_2}^{y_2' + 1} \left[ c_{ilj} + c_{il(j+1)} + \left( y_1^{il(j+1)} \right) + \left( y_1^{il(l+1)(j+1)} \right) + 16 \right] dy_2 
\end{align*}
\]
Recall that $\tilde{y}_1^{ilj} = a_{ilj} + b_{ilj}y_2$. Integrating the final integral over $y_2$ and plugging in the values of $a_{ilj}$ and $b_{ilj}$ gives us: 
\[
\begin{align*}
\frac{(y_4')}{2} \left( y_3' \right) \int_{y_2}^{y_2' + 1} \left[ a_{ilj} + a_{il(j+1)} + a_{il(l+1)(j+1)} + b_{ilj}y_2 + b_{il(j+1)} + b_{il(l+1)(j+1)} + 16 \right] dy_2 \\
= \frac{(y_4')}{2} \left( y_3' \right) \int_{y_2}^{y_2' + 1} \left[ \begin{array}{c}
\sum y_1^{il(j+1)} + y_1^{il(l+1)(j+1)} + \frac{y_1^{il}(j+1)}{y_1^{il} - y_1^{il(j+1)}} + \frac{y_1^{il(l+1)(j+1)}}{y_1^{il} - y_1^{il(l+1)(j+1)}} \\
\end{array} \right] dy_2 + 16 
\end{align*}
\]
Again, define $y_2' = y_2^{(i+1)} - y_2^l$, so this integral portion can be written: 
\[
\begin{align*}
\frac{(y_4')}{2} \left( y_3' \right) \frac{(y_2')}{2} \left[ \begin{array}{c}
\sum y_1^{il(j+1)} + y_1^{il(l+1)(j+1)} \\
\end{array} \right] 
\end{align*}
\]
The Probability of making a sale conditional on a given package as:
\[
\begin{align*}
P^j &= \left( \frac{1}{4096} \right) \sum_{i=1}^{4} \sum_{l=1}^{4} \sum_{j=1}^{4} \int_{y_2'}^{y_2' + 1} \int_{y_3'}^{y_3' + 1} \left[ e_{ilj} + f_{ilj}y_4 + 4 \right] dy_4 dy_3 dy_2 \\
&P^j &= \left( \frac{1}{4096} \right) \sum_{i=1}^{4} \sum_{l=1}^{4} \sum_{j=1}^{4} \left( \frac{(y_4')}{2} \left( y_3' \right) \left( y_2' \right) \right) \frac{(y_4')}{2} \left( y_3' \right) \frac{(y_2')}{2} \left[ \begin{array}{c}
\sum y_1^{il(j+1)} + y_1^{il(l+1)(j+1)} \\
\end{array} \right] 
\end{align*}
\]
I calculate \( P^j \) using the following algorithm:

1. (a) Begin with \( y_1^1 = y_3^1 = y_2^1 = -4 \). Initialize \( \mathbf{Y} = 0 \).
   
   (b) Evaluate the \( H(\theta) \) function at every node of \( y_1 \).

   (c) If there is a sign change at any of the nodes, find the value of \( y_1^{ij*} \) for the particular node of \( i, l \) and \( j \).

   (d) If there are no sign changes, move on to the next value of \( y_2 \).

   (e) Continue steps (b)-(c) for \( i = 1, 5, l = 1, 5 \) and then \( j = 1, 5 \) to calculate all the values of \( y_1^{ij*} \). Store this vector as vector \( y_1^* \) with dimension (125 x 1)

   • Note: If \( H(\theta) < 0 \) for all nodes of \( y_2^i, i = 1, 5 \) \( \mid y_4, y_3 \) then the probability of making a sale is zero for those values because the value of the current package is less than the reservation value.

   • Note: If \( H(\theta) > 0 \) for all nodes of \( y_2^i, i = 1, 5 \) \( \mid y_4 = y_3 = -4 \), then the probability of making a sale is one for those values because the value of the current package is greater than the reservation value.

   (f) Given the vector \( y_1^* \), cycle through the values of \( i, l, j \), picking off the appropriate \( y_1^* \) values and computing the appropriate \( y_4', y_3', \) and \( y_2' \) cell length.

   (g) Sum up the values as we cycle through and then divide the entire sum by 4096 (8^4) which gives us the probability of making a sale conditional on a package of type \( \mathbf{x} \), and a distribution of preferences, \( \Phi(\theta) \).

   (h) NOTE: There needs to be an adjustment made for the cases where \( \theta_1 = \beta \) is negative because consumers with that preference vector would not participate in the market.

**B Appendix: Data Construction**

This appendix provides additional detail on the construction of the number of encounters and transformation of county level median prison and probation sentence data.

First, consider the data I use to estimate the structural model parameters. One issue that arises is how to define a ‘retail’ purchase. Any package weight threshold is somewhat
arbitrarily determined and researchers have used a wide variety of thresholds values including 1, 3.5, 28 and 140 grams.\textsuperscript{55} I use 3 grams for the weight threshold defining a retail purchase because that weight is consistent with about ten use sessions. However, there is no strong consensus on what weight constitutes a retail amount of cocaine because there may exist both small and large users who routinely purchase very different weights.

B.1 Number of Encounters

In order to derive the probability of arrest for sale and possession of cocaine, I need an estimate of the number of encounters made by dealers and consumers. Recall equation (15) which calculates the total number of searches made by consumers to define the number of encounters \(E\).

For simplicity at this point I choose to estimate \(E\) outside of the model rather than use equation (15). I calculate the number of encounters as the number of cocaine consumers within the county \(M_{l}\) times the average number of purchases the average consumer makes, \(\pi\):

\[
E = \pi \times M_{l}
\]

The average number of purchases \(\pi\) is provided by the Office of National Drug Control Policy (ONDCP) and is taken from the Arrestee Drug Abuse Monitoring (ADAM) study.\textsuperscript{56}

\(M_{l}\) represents the number of cocaine consumers within county \(l\). This is calculated using data from the National Household Survey on Drug Abuse (NHSDA) and census data. First, I run a probit model to estimate the effect of demographic characteristics on the probability of consuming cocaine within the last month for each year of the NHSDA. The null hypothesis that probit estimates are stable across years is rejected at the 5% level. Table 9 presents the coefficient estimates of the probit model for the year 2001. The year-specific probit coefficient estimates are used with census data to calculate the probability the average citizen of county


\textsuperscript{56}See Riley (1997). It is important to note that \(\pi\) is provided by ADAM, which focuses on recent arrestees. This particular report provides information on drug purchase and use patterns in six select cities. Powder cocaine users report mean number of purchases in the previous week ranging from 3 to 9 (depending on the city), and crack cocaine users report the mean number purchases ranging from 6 to 14. Because this study oversamples drug offenders, I use data from arrestees charged with non-drug offenses to calculate an average purchase statistic.
would use cocaine within the last month. Multiplying by the population of county $l$ results in an estimate of the number of cocaine consumers $M_l^C$.

### B.2 Transformation of Penalty Data

The sentence length data comes from the National Judicial Reporting Program. The NJRP provides detailed information on the sentences that convicted felons receive in state courts on a biannual basis. Drug offenses are split into drug trafficking (which includes manufacturing, distributing, selling, smuggling, and possession with intent to sell as well as attempts to do any of the above), and drug possession (which includes possession of an illegal drug, attempts to possess an illegal drug, but excludes possession with intent to sell). The NJRP data includes type of sentence (prison, probation, community service, etc.) for each felon reported as well as the length of their particular sentence. In order to get an accurate measure of the sentence imposed for drug offenses, all felons with multiple offenses were dropped. This leaves observations of sentences of felons for whom the drug offense was the only crime of record.

Some values of the sentences were unusually large, and life sentences (and variations on life sentences) were coded with extremely high values. In order to create meaningful statistics, it was necessary to transform the way sentences were coded. Agents consider the sentence as a cost, and I allow dealers to discount the future value of months served. That is, agents do not weigh a month in prison served 20 years from now at the same cost as a current month served in prison. The prison sentences were transformed according to the function $S = \sum_{t=1}^{T} \beta^t$ where $T$ is the length of the sentence and $\beta$ is the discount factor. I used a discount factor of $\beta = 0.9$. Probation terms were transformed in a similar way. In some cases, dealers received a prison sentence and parole time after the sentence. I had to account for any time served in prison before beginning the parole term. The transformed probation function is $P = \sum_{t=T+K}^{T} \beta^t$ where $T$ is the length of the prison sentence, $K$ is the total length of prison sentence plus parole term, and $\beta$ is the discount factor. Once the variables were transformed, I calculated the median of all the sentences for a given county-time combination.
B.2.1 County-Year Combinations

The complete data set matches county level prison sentences (and probation terms) and county level arrest rates to the purchase data (matched by FIPS code and year). However, a good deal of the purchase data comes from counties that are not represented in the NJRP or UCR. NJRP data is only collected every two years with data on drug sale and possession available from 1986-1998. If the mean penalty was calculated for each county-year and matched to the drug purchase data, the total number of matched observations would be small.\(^{57}\) For this reason, I assume the values of median sentence length is stable and assign the overall county median to all years.\(^{58}\)

However, prior to assigning the missing county-year values, I test this assumption.

B.2.2 Hypothesis Test

First, let me introduce some notation designed for this hypothesis test only (\(i.e.,\) the \(y\) used in this section is not associated with the \(y\)'s used in previous appendices). Let

\[
y_{ijt} = \delta_{it} + \varepsilon_{ijt} \quad (35)
\]

\(j = 1, 2, ..., J\)

\(i = 1, 2, ..., I\)

\(t = 1, 2, ..., T\)

where \(y_{ijt}\) is the \(j^{th}\) sentence in the \(i^{th}\) county at time \(t\). The median for county \(i\) at time \(t\) is given by \(\delta_{it}\) and the error \(\varepsilon_{ijt}\) is assumed to have some unknown distribution \(F(0, \sigma)\). There are \(n\) total observations of sentences.

I wish to test the hypothesis that county behavior does not change over time:

\[
H_0 : \delta_{i1} = \delta_{i2} = ... = \delta_{iT} \quad (36)
\]

\(H_A : \delta_{i1} \neq \delta_{i2} \neq ... \neq \delta_{iT}\)

\(\forall i = 1, ..., I\)

\(^{57}\)When the penalty and purchase data are merged on both year and county, there are a total of 1995 observations of purchases with weight less than 3 grams.

\(^{58}\)I use median sentence length as opposed to mean sentence due to the coding of the data. Life sentences, and variations on life sentences are coded with extremely large values which will skew the mean. The median will not give undue weight to the outliers.
If the null is rejected, then I cannot assign the county median to all years of purchase data within that county.

The test statistic for this hypothesis test is:

\[ z_i = (A_i \hat{\delta}_i)'[A_i D_i(\delta) A_i']^{-1}(A_i \hat{\delta}_i) \sim \chi^2_{(n_i)} \]  

(37)

where \( \hat{\delta}_i \) is the vector of estimated year medians for county \( i \), and \( D_i(\delta) \) is the \( n_i \times n_i \) diagonal covariance matrix of the medians. \( A_i \) is an \( (n_i - 1) \times (n_i) \) matrix where \( n_i \) is the number of years of available data in county \( i \). For instance, if there were 5 years of penalty data available for county \( i \), then \( A_i \) would take the form:

\[
A_i = \begin{bmatrix}
1 & -1 & \ldots & \ldots & \\
\ldots & 1 & -1 & \ldots & \\
\ldots & \ldots & 1 & -1 & \\
\ldots & \ldots & \ldots & 1 & -1 \\
\end{bmatrix}
\]

(38)

The test statistic \( z_i \) is distributed chi-squared with \( (n_i) \) degrees of freedom.\(^{59}\)

In order to find the diagonal terms of \( D_i(\delta) \), I need to find the variance of each county-time median. Following Bushinsky (1997), the estimate of the variance of a median is:

\[
\hat{\sigma}_{it} = \frac{1}{4n_{it} \times f^2(0)} 
\]

(39)

where \( n_{it} \) is the number of observations in county-time combination \( i-t \), and \( f(0) \) is the density of \( \varepsilon \) evaluated at 0.

However, this density \( f \) is unknown. Rather than make any assumptions regarding the distribution of \( \varepsilon \), I estimate \( f(0) \) using a kernel estimator. Following the Rosenblatt-Parzen kernel estimator in Pagan and Ullah (1999) the kernel estimator is:

\[
\hat{f}(0) = \frac{1}{n \times h} \sum_{l} K \left( \frac{\varepsilon_l}{h} \right) 
\]

(40)

where \( K \) is the standard normal kernel, \( K(\psi) = (2\pi)^{\frac{-1}{2}} \exp(-0.5\psi^2) \), \( n \) is the sample size, and \( h \) is the bandwidth which is a function of the sample size and goes to 0 as \( n \to \infty \).

\(^{59}\)The degrees of freedom will depend on the dimensions of the covariance matrix. For the possession data, this is 961 degrees of freedom. 961 is the total number of distinct county-time combinations (several missing counties and years). For the sale data, the degrees of freedom is 1635.
I find the bandwidth that minimizes the AMISE (approximate mean integrated squared error) is $h = 1.0417$ and $n = 90356$ for possession sentence data. The bandwidth for sale data is $h = 1.9481$ and $n = 123973$. I can calculate $f(0)$, and $f^2(0)$.

\[
\begin{align*}
\text{Possession} & : f(0) = 0.08555 \\
\text{Sale} & : f(0) = 0.05250
\end{align*}
\]

From this point, I calculate a vector of variances using formula (39). Once I calculate $\tilde{\sigma}_{it}$, I construct the $D_i(\delta)$ matrix, and construct the test statistic $z_i$ for each county. The value of the critical value $t$ varies over counties due to variation in the number of years available for a given county. The null is not rejected for any county in the penalty data. Therefore, county penalty behavior is stable over time, and I assign the county median sentence to all purchase observations made in that county.
**Figure 1: Increased Enforcement (Cost) in the Illegal Drug Market**
(Affecting Supply Only)

Note: This approach excludes any response to enforcement by consumers.

**Figure 2: Increased Enforcement in the Illegal Drug Market**
(Affecting both Supply and Demand)

Note: Equilibrium approach including demand response.
Figure 3: Finding the Reservation Value

Table 1: Summary Statistics of Merged STRIDE and NJRP Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (grams)</td>
<td>0.667</td>
<td>0.737</td>
</tr>
<tr>
<td>Purity</td>
<td>0.740</td>
<td>0.210</td>
</tr>
<tr>
<td>Price ($)</td>
<td>82.362</td>
<td>162.175</td>
</tr>
<tr>
<td>Price/Pure gram</td>
<td>272.449</td>
<td>955.046</td>
</tr>
<tr>
<td>Possession Probation</td>
<td>0.167</td>
<td>0.611</td>
</tr>
<tr>
<td>Possession Sentence</td>
<td>6.514</td>
<td>2.041</td>
</tr>
<tr>
<td>Sale Probation</td>
<td>0.039</td>
<td>0.262</td>
</tr>
<tr>
<td>Sale Sentence</td>
<td>9.698</td>
<td>0.851</td>
</tr>
</tbody>
</table>

Note: The number of observations is 11866. The penalty variables in the merged dataset have been transformed according to the methods described in Appendix B. We must be careful when discussing the means and variances of these variables, because they are not easily translated into years or months.
Table 2: Summary Statistics of NJRP Data Alone

<table>
<thead>
<tr>
<th>Sentence Length (Years)</th>
<th>Percent Of Sample</th>
<th>Mean (Years)</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sale:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Sentence</td>
<td>5.9</td>
<td>Life</td>
<td>N/A</td>
</tr>
<tr>
<td>0-5 Years</td>
<td>70.8</td>
<td>1.55</td>
<td>1.27</td>
</tr>
<tr>
<td>5-20 Years</td>
<td>12.7</td>
<td>8.83</td>
<td>3.05</td>
</tr>
<tr>
<td>Possession:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Sentence</td>
<td>2.2</td>
<td>Life</td>
<td>N/A</td>
</tr>
<tr>
<td>0-5 Years</td>
<td>87.3</td>
<td>1.14</td>
<td>0.999</td>
</tr>
<tr>
<td>5-20 Years</td>
<td>4.9</td>
<td>9.25</td>
<td>2.85</td>
</tr>
</tbody>
</table>

Note: Sentences between 20 years and Life are not reported in this Table.

Table 3: Yearly Average Number of Arrests per County Conditional on Positive Arrests for Possession and Sale of Cocaine/Opium

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Counties Reporting</th>
<th>Average Arrests for Cocaine/Opium Sale</th>
<th>Average Arrests for Cocaine/Opium Possession</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>972</td>
<td>59.34</td>
<td>169.39</td>
</tr>
<tr>
<td>1986</td>
<td>1140</td>
<td>80.61</td>
<td>176.58</td>
</tr>
<tr>
<td>1987</td>
<td>1115</td>
<td>97.47</td>
<td>216.29</td>
</tr>
<tr>
<td>1988</td>
<td>1279</td>
<td>120.41</td>
<td>244.00</td>
</tr>
<tr>
<td>1989</td>
<td>1512</td>
<td>136.70</td>
<td>247.32</td>
</tr>
<tr>
<td>1990</td>
<td>1463</td>
<td>142.91</td>
<td>218.83</td>
</tr>
<tr>
<td>1991</td>
<td>1417</td>
<td>139.64</td>
<td>201.21</td>
</tr>
<tr>
<td>1992</td>
<td>1449</td>
<td>132.94</td>
<td>210.68</td>
</tr>
<tr>
<td>1993</td>
<td>1315</td>
<td>139.67</td>
<td>224.31</td>
</tr>
<tr>
<td>1994</td>
<td>1775</td>
<td>111.64</td>
<td>196.54</td>
</tr>
<tr>
<td>1995</td>
<td>1878</td>
<td>101.61</td>
<td>185.15</td>
</tr>
<tr>
<td>1996</td>
<td>1826</td>
<td>92.46</td>
<td>169.47</td>
</tr>
<tr>
<td>1997</td>
<td>1854</td>
<td>89.03</td>
<td>177.75</td>
</tr>
<tr>
<td>1998</td>
<td>1830</td>
<td>91.96</td>
<td>182.52</td>
</tr>
<tr>
<td>1999</td>
<td>1787</td>
<td>86.12</td>
<td>171.63</td>
</tr>
<tr>
<td>2000</td>
<td>1830</td>
<td>76.22</td>
<td>168.30</td>
</tr>
</tbody>
</table>
Table 4: Relationship between Pure Gram Price and Expected Enforcement Variables

<table>
<thead>
<tr>
<th>Agent</th>
<th>Enforcement Variable</th>
<th>Estimate (1)</th>
<th>Estimate (2)</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer:</td>
<td>Expected Probation Term (Possession):</td>
<td>-147.887*</td>
<td>-1.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Expected Prison Sentence (Possession)</td>
<td>2.868</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intercept</td>
<td>229.580*</td>
<td>5.35</td>
<td></td>
</tr>
<tr>
<td>Dealer:</td>
<td>Expected Probation Term (Sale)</td>
<td>-238.962</td>
<td>-0.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Expected Prison Sentence (Sale)</td>
<td>-18.105</td>
<td>-0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intercept</td>
<td>235.540*</td>
<td>5.55</td>
<td></td>
</tr>
</tbody>
</table>

Note: Only year fixed effects are included in this regression. Standard Errors are in parentheses.

Table 5: Relationship between Pure Gram Price and Arrest Rates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrest Rate Sale</td>
<td>-179.73</td>
<td>195.09</td>
<td>-0.92</td>
</tr>
<tr>
<td>Arrest Rate Possession</td>
<td>-244.40</td>
<td>119.11</td>
<td>-2.05</td>
</tr>
</tbody>
</table>

Note: Only year fixed effects are included in this regression.
Table 6: Relationship between Purity and Expected Enforcement Variables

<table>
<thead>
<tr>
<th>Agent</th>
<th>Variable</th>
<th>Estimate (1)</th>
<th>Estimate (2)</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer</td>
<td>Expected Sentence (Possession)</td>
<td>0.020*</td>
<td>(0.004)</td>
<td>4.83</td>
</tr>
<tr>
<td></td>
<td>Expected Probation (Possession)</td>
<td>0.094*</td>
<td>(0.015)</td>
<td>6.09</td>
</tr>
<tr>
<td></td>
<td>Intercept</td>
<td>0.662*</td>
<td>(0.009)</td>
<td>74.51</td>
</tr>
<tr>
<td>Dealer</td>
<td>Expected Sentence (Sale)</td>
<td>0.021*</td>
<td>(0.004)</td>
<td>5.20</td>
</tr>
<tr>
<td></td>
<td>Expected Probation (Sale)</td>
<td>0.223*</td>
<td>(0.100)</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td>Intercept</td>
<td>0.665*</td>
<td>(0.009)</td>
<td>75.56</td>
</tr>
</tbody>
</table>

Note: Standard Errors in parentheses.

Table 7: Correlation Matrix of Expected Enforcement Variables

<table>
<thead>
<tr>
<th></th>
<th>Expected Sentence (Sale)</th>
<th>Expected Sentence (Possession)</th>
<th>Expected Probation (Sale)</th>
<th>Expected Probation (Possession)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Sentence (Sale)</td>
<td>1.0000</td>
<td>0.3756</td>
<td>-0.0382</td>
<td>-0.2198</td>
</tr>
<tr>
<td>Expected Sentence (Possession)</td>
<td>0.3756</td>
<td>1.0000</td>
<td>-0.0215</td>
<td>0.4930</td>
</tr>
<tr>
<td>Expected Probation (Sale)</td>
<td>-0.0382</td>
<td>-0.0215</td>
<td>1.0000</td>
<td>0.1991</td>
</tr>
<tr>
<td>Expected Probation (Possession)</td>
<td>-0.2198</td>
<td>0.4930</td>
<td>0.1991</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
Table 8: Correlation Matrix of Enforcement Variables

<table>
<thead>
<tr>
<th></th>
<th>Median Sentence (Sale)</th>
<th>Median Sentence (Possession)</th>
<th>Median Probation (Sale)</th>
<th>Median Probation (Possession)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Sentence (Sale)</td>
<td>1.0000</td>
<td>0.8333</td>
<td>-0.3903</td>
<td>-0.1173</td>
</tr>
<tr>
<td>Median Sentence (Possession)</td>
<td>0.8333</td>
<td>1.0000</td>
<td>-0.1666</td>
<td>-0.1353</td>
</tr>
<tr>
<td>Median Probation (Sale)</td>
<td>-0.3903</td>
<td>-0.1666</td>
<td>1.0000</td>
<td>0.1315</td>
</tr>
<tr>
<td>Median Probation (Possession)</td>
<td>-0.1173</td>
<td>-0.1353</td>
<td>0.1315</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Note: These enforcement variables do not account for the probability of arrest for a particular drug offense.

Table 9: Probit Regression Results – NHSDA 2001

<table>
<thead>
<tr>
<th></th>
<th>Coefficient Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.95935</td>
<td>0.00232</td>
</tr>
<tr>
<td>Male</td>
<td>-0.22480</td>
<td>0.00122</td>
</tr>
<tr>
<td>Never Married</td>
<td>-0.65953</td>
<td>0.00221</td>
</tr>
<tr>
<td>Black</td>
<td>0.26882</td>
<td>0.00391</td>
</tr>
<tr>
<td>Age 16-20</td>
<td>-0.02368</td>
<td>0.00153</td>
</tr>
<tr>
<td>Age 35-49</td>
<td>-0.23400</td>
<td>0.00323</td>
</tr>
<tr>
<td>HS Grad or GED</td>
<td>-0.24664</td>
<td>0.00234</td>
</tr>
<tr>
<td>Population&gt;1million</td>
<td>-0.02953</td>
<td>0.02379</td>
</tr>
</tbody>
</table>

Note: Probit model results are presented for the year 2001 only. The null hypothesis that behavior is stable over time is rejected at the 5% level.