Submitted to *Management Science*
manuscript

Investigating Dynamic Effects of Counterfeits with a Hierarchical Random Changepoints Simultaneous Equation Model

Yi Qian
Department of Marketing, Kellogg School of Management, Northwestern University, Evanston, IL 60208, yiqian@northwestern.edu

Hui Xie
Division of Epidemiology & Biostatistics, University of Illinois, Chicago, IL 55599, huixie@uic.edu

With growing emerging markets and globalization, counterfeiting has gained pervasiveness and presented worldwide impacts. Using a unique panel dataset and a new model, this article aims at enhancing the understanding of brand management against counterfeits. Our model extends two important business analytic tools, Hierarchical Bayesian (HB) and Random Changepoints models, to simultaneously take into account the following important data features when studying firms' responses to counterfeit entry: (1) *Endogeneity* of counterfeit entry as well as the moderators of the entry effects, (2) *Unobserved heterogeneity* in both magnitudes and timing of firm responses, (3) *Discontinuous* changes in response to counterfeit entry, and (4) *regime-switching* moderating effects. The proposed methodology improves the estimation of firms' strategies with heterogeneous response times, and substantially increases the power to identify firm attributes that moderate the competitive effects. We identify both a temporary negative short-term effect and a stable positive long-term effect of counterfeit sales on the authentic prices. The finding of dynamic effects precisely unifies two strands of I.O. theories on the pricing impacts of competition. Such dynamic effects were not identified in a standard IV model that ignores heterogeneous changepoints. The proposed model allows us to better identify what moderates the dynamic effects. Because our model extends the popular HB models to allow for regime-switching moderating effects with unit-specific latent regime changepoints, it substantially improves the power to detect moderating effects, some of which could not be detected in conventional HB models. We account for the potential endogeneity of moderating variables through a latent instrumental variables (LIV) approach. The improved estimation enhances our understanding of firms' responses to counterfeit entry in emerging markets. The hierarchical dynamic effect analysis reveals that (1) pre-entry product quality *moderates* the short-term price competition effects; (2) brand popularity *moderates* the long-term price increase effect, and (3) firms with more innovation, less diversification from infringed markets or more human capital were faster in responding to and differentiating from counterfeits.

Key words: Multiple Changepoints; Counterfeit; Hierarchical Bayesian; Intellectual Property Right; Management Strategy; Response Time.

1. Introduction

With the fast growth in emerging markets, consumption of branded products rapidly gain share in these markets and subsequently in the world economy. Parallel to this trend is the increasing presence of counterfeits. The Federal Bureau of Investigation reports that U.S. companies lose $200-250 billion annually due to worldwide copyright, trademark, and trade-secret infringements. The European Commission (EC) reckons that the value of counterfeiting as a percentage of world trade is growing. Between 1990 and 1999, it doubled from 3.5 percent to 7 percent (Choate 2005). A deep understanding of brand management in the face of counterfeiting in these emerging markets is critical for academics and practitioners. The entry of counterfeiters can have two opposite impacts on the authentic producers. On the one hand, their entry potentially exerts competitive pressure on authentic firms. Authentic prices could also drop as a result of limit (predatory) pricing strategy (Carlton and Perloff, 2005). On the other hand, their entry may lead to increases in authentic prices due to the segmentation of price-sensitive and insensitive consumers in the market (Frank and Salkever 1997), or due to authentic producers' innovations and self-differentiation mechanics to alleviate competition (Qian 2008). The former competitive effect is likely to take place immediately upon entry, while the latter effect may arrive with some lags. This lag can differ from company to company due to inherent firm heterogeneity in their ability to respond to market shakeups caused by counterfeit entry.

Therefore, *how much*, in *which direction*, *when* and *why* authentic firms' marketing norms (e.g. prices) change in response to counterfeit entry are both interesting and pertinent questions to address. It is also of substantial interest to identify drivers that explain the potential inter-firm differences in their response behaviors. When studying these questions, it is worth noting several important empirical identification challenges. First, counterfeit entry is unlikely to be exogenous. Counterfeiters are more likely to infringe upon a brand if the authentic product is easier to imitate, has a larger markup, or if the brand management is worse. Under such circumstances, counterfeit entry will be correlated with authentic prices. However, a causal link cannot be inferred from this correlation. The second challenge is discontinuous structural changes in model parameter values. As noted above, it is natural for the authentic firms to take time to analyze the changing business environment, design corresponding strategies, and implement them in practice. The stable long-term effect follows only with some delay in time, which we refer to as response time. Such delay in firm (or consumer) response behaviors has long been recognized (e.g., Kotler 1971, Robinson 1988, Bowman and Gatignon 1995). In these scenarios, the empirical model parameters capturing the effect of a market change do not remain unchanged in that an authentic firm changes its state underlying the marketing outcome at some changepoints. This phenomenon is likely to be more salient in emerging markets with immature infrastructure from which our data come from: in these market environments, firms are more likely to adopt different short-term and long-term strategies in the face of competitive shocks. Hence it is crucial to incorporate and identify any *discontinuous* changes in firms' response behaviors. The third challenge is the unobserved heterogeneity in individual firms' response magnitude and timing. In particular, it is often the case that authentic firms respond to a market change with different response times (e.g., Robinson 1988, Bowman and Gatignon 1995). Firms in emerging markets often have diverse background and hence exhibit significant *heterogeneous* pattern and timing in their competitive strategies. It is important to account for the heterogeneity in latent response timing in order to obtain accurate ascertainment of counterfeit entry effects. The forth challenge is to account for two potentially important issues in assessing drivers of heterogeneity in firms' response behaviors: the potential regime-switching moderating effects (i.e., moderating effects varies across regimes (sample periods) separated at unknown unit-specific changepoints) and the potential endogeneity of moderating variables that affect firms' response behaviors. These issues are typically ignored in commonly-used Hierarchical Bayesian (HB) models but can introduce substantial bias.

We propose here a hierarchical random-changepoints simultaneous equation model to address these empirical challenges. Our framework builds on and extends the ideas from the following literatures important for business analytics: random changepoints (RC) models for structural changes, HB models for heterogeneity, and IV methods for endogeneity. In terms of the methodological contribution, our modeling framework extends two popular/emerging business analytics tools (RC and HB models) in the following important ways: (1) We extend the well-established random changepoints model, an emerging business analytic tool, to allow for endogenous regressors, thereby enabling *automatic* and *cleaner* assessment of *discontinuous* changes in response to *endogenous* counterfeit entry; (2) Unlike the traditional HB model, our hierarchical changepoints model allows the moderators of counterfeiting to have regime-varying moderating effects separated at latent unit-specific regime changepoints, and consequently improves the power to detect moderators of counterfeiting impacts; (3) Unlike the standard HB model, we apply the latent IV technique (Ebbes, Wedel, Böckenholt, and Steerneman 2005, Zhang, Wedel and Pieters 2009) to account for the potentially endogenous moderators of counterfeiting effects; (4) It extends the conventional HB model to include latent response time, and consequently this new approach enables us to study what affect firms' response time to counterfeit entry, an important dimension in firms' response behavior; and (5) Our Bayesian approach provides a tractable and coherent framework to account for all these issues. The new methodology enables better understanding of firms' dynamic response behaviors under heterogeneous response times, minimize bias in the entry effect estimates, and substantially increase the power to identify firm attributes that moderate these effects. In observational studies such as ours, the above important issues (endogeneity, discontinuous structural changes, and unobserved heterogeneity) are frequently encountered, and therefore these extensions have the potential to have wide implications.

Substantively, the benefits of our modeling approach enhance our understanding of counterfeiting issues in emerging markets. First, the new modeling and testing approach are capable of depicting a more complete picture of the impacts of counterfeits, and provide convincing statistical evidence for the short-term negative and long-term positive pricing effects of counterfeit entry. [1] By allowing for unobserved heterogeneity in firms' latent response times and thus more effectively disentangling these different stages of impacts using the proposed model, our analysis reveals dynamic and stronger effects of counterfeit entry. The new finding on the dynamic effects precisely unify two strands of I.O. theories on the pricing effects of competition. We further investigate drivers that can explain the inter-firm differences in their response behaviors, an interesting topic not yet studied in prior research. Using the improved dynamic effect estimates and other extensions of HB models as described above, our analysis reveals several interesting findings. The hierarchical dynamic effect analysis finds that the pre-entry authentic product quality *moderates* the short-term price competition effects of counterfeit entry and is helpful in alleviating the harmful impacts of counterfeit entry. Furthermore, brand popularity *moderates* the long-term price increase effect, and firms that were less popular pre-entry tend to have more price increase, consistent with the hypothesis that counterfeits can serve as free-advertising for their authentic counterparts. Our hierarchical change-point analysis shows that firms with more innovation, less diversification from infringed markets or more human capital were faster in responding and differentiating from counterfeits.

2. Modeling Approach

In this section we develop the proposed methodology to study the effects of counterfeit entry and sales on the authentic product prices using a dataset on the Chinese shoe industry. The panel data consist of annual average prices, costs and sales for 31 authentic branded companies and their counterfeits from the year 1993 to 2004 (Qian 2008). As a preliminary analysis, Figure 1 presents the time plot of the average log deflated authentic high-end product prices. Specifically the figure plots the regression coefficients on a set of dummies indicating the number of years relative to the year of counterfeit entry with the log deflated authentic price as the response variable. The plot suggests the presence of a discontinuous change in effects of counterfeit entry on authentic prices: there was a reduction in the average authentic prices within first two years of the counterfeit entry, after which there was an increase in the average authentic prices.

[1] Such dynamic counterfeit entry effect is not tested in related prior research (Qian 2008). The standard econometric technique employed in Qian (2008) addressed her research question of the average treatment effects of counterfeiting but was not able to detect a statistically significant short-term negative effect of counterfeit entry. Our new analyses reveals dynamic and stronger effects of counterfeit entry, which are also crucial for analyzing important moderator effects.

The above simple analysis, though informative and useful, has some important limitations. Notably, the analysis has not yet accounted for the potential endogeneity issue of counterfeit entry. Secondly, although the analysis reveals a potential changepoint in the average price profile, it ignores the heterogeneity of the changepoint among the firms, and assumes that all the authentic firms took the same amount of time to respond to their counterfeit entries. As aforementioned and further demonstrated in the following sections, ignoring the heterogeneity can attenuate the effect estimates of counterfeit entry. Bias also arises when assessing factors moderating these effects. Furthermore, the preliminary analysis does not allow us to study what affects a firm's response time to counterfeit entry.

To overcome the limitations of the preliminary analysis, we propose a hierarchical random-changepoints simultaneous equations model to investigate the effects of counterfeit entry. Our modeling framework jointly models the panel price profiles of the authentic firms evolved over time, the quantity of counterfeits faced by the authentic firms as well as the latent random changepoints in the panel outcome profiles. Below we describe the overall model.

2.1. Within-firm Model

Let $Y_1, ..., Y_N$ denote the outcome vectors on a random sample of N units (i.e. firms in our application), where $Y_i = (Y_{i1}, ..., Y_{iT})$ is a T-dimensional panel outcome vector for the ith unit, $i = 1, ..., N$ and $N = 31, T = 12$ in our application. We seek to develop a modeling approach that simultaneously accounts for three important empirical features: (1) Discontinuous changes in effects of a market change; (2) Endogeneity of the market change and moderating variables; and (3) Heterogeneity across market agents in the magnitude of changes, and the heterogeneity of *timing* at which different stages of effects kicked in.

We first describe the modeling strategy for effect instability. One approach is to use lagged covariate values through the distributed lag regression models (Almon 1965), where the response variable is specified as a function of current and past covariate values as follows

$$Y_{it} = \sum_{l=0}^{L} X'_{i,t-l}\beta_l + U'_{it}\alpha_i + W'_{it}\gamma + \epsilon^Y_{it},$$

where X refers to the independent variable of primary interest; the index $t = 1, \cdots, T$; L is the lag length, and $\beta_l, l = 1, \cdots, L$, denote the lagged effects. The variable, U, includes those whose effects on Y are heterogeneous among firms. This may include, though not limited to, the unit dummy variables to capture the time-constant unobserved heterogeneity across units. The variable, W, may include the time dummy variables to capture the common market shocks to all the units at a given time. To avoid the common multicollinearity issue among the lagged variables, a distributed lagged function is imposed on the set of parameters $(\beta_0, \beta_1, \cdots, \beta_L)$. However, prior information on response time is rarely available and thus it can be difficult to specify a sensible lag function. Furthermore this approach provides no basis to study the firms' response time.

Our approach is to adopt the random-changepoints modeling technique to allow for multi-stage entry effects. The random-changepoints models are well-established modeling techniques for structural changes and have been widely used for modeling parameter instability with early applications in statistics and economics (e.g. Barry and Hartigan 1993, Carlin et al. 1992, Lange et al. 1992, Bai 1997, Bai and Perron 1998, Chib 1998). More recently, Fader et al. (2004), Fong and Desarbo (2007) and Schweidel and Fader (2009) extend the random-changepoints models to business applications. In our application, we posit the following random-changepoints model for within-firm responses

$$Y_{it} = X_{it}'\beta_{ik} + U_{it}'\alpha_i + W_{it}'\gamma + \epsilon_{it}^Y, \qquad t = T_{ik}, \cdots, T_{i(k+1)} - 1, \;\; k = 0, \cdots, K. \tag{1}$$

In the random-changepoints model, the covariate X includes variables that are believed to have differential multiple-stage effects; K denotes the number of changepoints which partition the time series to $K+1$ regimes; T_{ik} and $T_{i(k+1)}$ denote the two endpoints for the kth stage with their values being unobserved and inferred from the outcome trajectory; and notation-wise we set $T_{i0} = 1 \; \forall i$. The random-changepoints model is flexible, and nests some interesting models as special cases. Notably, when each time point is a changepoint ($K = T - 1$), it becomes the most general model

$$Y_{it} = X_{it}'\beta_{it} + U_{it}'\alpha_i + W_{it}'\gamma + \epsilon_{it}^Y, \tag{2}$$

where β_{it} is allowed to vary over t in an arbitrary form. On the other extreme, when there is no changepoint ($K = 0$), it reduces to a static panel data model which assumes that the effect of X remains constant over time (i.e., $\beta_{it} = \beta_i, \; \forall t$). When $1 \le K < T - 1$, it is a model intermediate between these two extreme cases. When $K = 1$, X is said to have both short-term (β_{i0}) and long-term (β_{i1}) impacts on the outcome. Regarding the nature of the change, the model makes no restriction that derivative must exist at the time of change, and allows for sudden discontinuous changes in parameter values. Therefore random changepoints models are natural for modeling discrete changes in parameter values. Such regime changes are natural outcomes of some market variations, such as new competitive entries or technological innovations (Fader, Hardie and Huang 2004, Fong and DeSarbo 2007). As such, this approach is especially suited for our application, where the changes are sudden and discontinuous. When the changes are continuous, the random-changepoints model uses a piecewise constant model with flexible data-driven time intervals to approximate the smooth function. In addition to the modeling flexibility and interpretability, the random-changepoints model explicitly models the times where each different stage kicks in, and allows us to investigate the response timing, a feature naturally modeled in changepoints models. We therefore build our approach from the random-changepoints models.

In the marketing area, Fader et al. (2004) develop a dynamic changepoint model that allows the changepoint process itself to evolve over time. They show that the dynamic changepoint model improves the new product sales forecasting. The method has been further studied and extended by Schweidel and Fader (2009). Fong and DeSarbo (2007) study Bayesian variable selection problems in multiple regression models with changepoints. Our approach is more similar to that of Fader et al. (2004) in that we explicitly model the random changepoints which are heterogeneous among the study units in a panel data setting. However, there are several important differences. First, unlike their study which focuses on forecasting new product sales, the focus of our study is on an accurate ascertainment of causal effects of a market change. As such, we must explicitly deal with the endogeneity issue, an important issue in empirical studies. To the best of our knowledge, the endogeneity issue in a hierarchical random-changepoints model has not been dealt with previously, although its importance has been discussed in Fader et al. (2004).[2] Furthermore, our MCMC algorithm avoids evaluating the model likelihood as required in the MLE approach of Fader et al. (2004). As a result, computationally it is more scalable to high-dimensional multiple changepoints problems. Last, the hierarchical structure in our model allows for incorporating covariates into a model for changepoints.

Related to the random changepoints models are the hidden Markov models. Böckenholt and Dillon (1997), Poulsen (1990), Ramaswamy (1997) and Netzer et al. (2008) use the hidden Markov Models to study the changes in segment memberships over time, which can be viewed as a special case of more general changepoints models (Fader et al. 2004, Fong and DeSarbo 2007). As will be shown later, the random changepoints model captures the cross-sectional heterogeneity of entry effects through a continuous representation of β_{ik} values. This is in contrast with the hidden Markov models where the cross-sectional heterogeneity is captured by a finite number of state variables. Past studies have demonstrated that continuous representation of heterogeneity is preferred (Allenby and Rossi 1999), by offering a more thorough control of agents' heterogeneity.

When the dynamics are assumed to be smooth, an alternative class of modeling approach is a state-space modeling, where the time-varying parameter is assumed to follow an autoregressive form that evolves over time smoothly. Neelamegham and Chintagunta (2004), Van Heerde, Mela and Manchanda (2004) and Lachaab et al. (2006) extended the Dynamic Linear state-space models to marketing applications. We choose to build on the random-changepoints model for our application because of (1) its suitability for modeling discrete changes as described above;[3] (2) its convenience

[2] A concurrent working paper by Perron and Yamamoto (2011) considers multiple structural changes in linear regression models with endogenous regressors. However as discussed in Section 3, our application is more complicated and our approach avoids several important difficulties encountered when extending their approach to panel data.

[3] For example, as pointed out in Netzer et al. (2008), such an continuous state space modeling approach is "inadequate to capture dynamics that are postulated to develop in a discrete manner such as an instantaneous regime shift in the market conditions or consumer preferences...". Lachaab et al. (2006) also noted that "Instead of a state space approach, a hidden Markov or change-point framework could be adopted as an alternative to accommodate discrete changes and structural shifts in the parameters."

in modeling firms' response time and factors affecting response time, an important dimension of firms' response behavior and (3) its convenience in modeling regime-switching moderating effects as will be described later.

In our modeling framework, we allow X to be endogenous. In our application, X_{it} refers to the time-varying quantity of counterfeit products faced by the authentic brand i. We follow the approach of Amemiya (1985) and introduce a latent variable X_{it}^* as follows

$$X_{it} = \begin{cases} X_{it}^* & \text{if } X_{it}^* \geq 0 \\ 0 & \text{if } X_{it}^* < 0 \end{cases} \tag{3}$$

$$X_{it}^* = \delta Z_{it} + \epsilon_{it}^X, \tag{4}$$

where X_{it}^* is a latent variable that determines the observed variable X_{it} according to Equation (3), and Z_{it} is a vector of exogenous instrumental variables that relate to X_{it}^*. The endogeneity of X_{it} is modeled by the correlation between the error terms ϵ_{it}^X and ϵ_{it}^Y, which are assumed to follow a bivariate normal:

$$\left[\epsilon_{it}^Y, \epsilon_{it}^X\right] \sim N(0, \Sigma_\epsilon), \;\; \Sigma_\epsilon = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}. \tag{5}$$

In this model, X is endogenous when σ_{12} is nonzero.

In the above simultaneous equation model, we are primarily interested in measuring the causal effects of changing X on the outcome Y using panel data where at least some of the units in the sample experienced the change of X value over the period under examination. The model allows multi-stage effects of changing X on Y: $\beta_i = (\beta_{i0}, \cdots, \beta_{iK})$ which denote a vector of $K+1$ parameters capturing $K+1$ stage effect of X on Y. Let T_{ik} be the latent time point at which the kth stage kicks in for the ith firm, and set $T_{i0} = 1$. Let $\tau_{ik} = T_{ik} - T_{i(k-1)}$ for $k > 1$, and $\tau_{i1} = T_{i1} - T_{ie}$ where T_{ie} is the time of counterfeit entry for the ith firm. Then τ_{ik} denotes the time for the ith firm to move from the $(k-1)$th stage to the kth stage. We note here the one-to-one correspondence between the latent variables $\tau_i = (\tau_{i1}, \cdots, \tau_{iK})$ and $T_i = (T_{i1}, \cdots, T_{iK})$. An ad-hoc method that specifies a common response time for firms, say τ_c, ignores this heterogeneity and can lead to biased estimates of the nonstationary effects of X on Y, as will be shown later. Our model aims to separate out these time-heterogeneous effects more cleanly by explicitly modeling the underlying unit-specific changepoints. Another benefit of doing this is to provide estimates of response times and to be able to study what affects firms' response times.

To model the response time τ_i as a function of firm-characteristics, we assume that there are continuous response time variables, τ_i^*, for the firm i. Because the observed value is determined by the coarsened units of the outcome, e.g. years in our dataset, the response times in τ_i^* are observed to fall in certain intervals with their actual values unobserved. Such data often occur in empirical studies, such as the length of time with the present employer or duration of unemployment in survey studies. In our case, the grouped value τ_i is determined by its underlying value τ_i^* according to the following set of rules:

$$\tau_{ik} = \begin{cases} 0 & \text{if } \tau_{ik}^* \leq 0 \\ m & \text{if } m-1 < \tau_{ik}^* \leq m, \;\; m > 0, \end{cases} \tag{6}$$

where m takes integer values. In the above model, $\tau_{ik} = 0$ implies that the firm does not experience $(k-1)$th stage, and the kth stage effect kicks in immediately for this firm.

We derive the likelihood for the above unit-level model as follows.

$$\begin{aligned} f(Y_i, X_i|\tau_i^*, \beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon) &= \int \sum_{\tau_i} f(Y_i, X_i, X_i^*, \tau_i|\tau_i^*, \beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon) dX_i^* \\ &= \int \sum_{\tau_i} f(Y_i, X_i, X_i^*|\tau_i, \beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon) f(\tau_i|\tau_i^*) dX_i^*, \end{aligned} \tag{7}$$

where the density function $f(\tau_i|\tau_i^*) = \prod_{k=1}^K f(\tau_{ik}|\tau_i^*)$, and $f(\tau_{ik}|\tau_i^*)$ is

$$f(\tau_{ik}|\tau_i^*) \propto \begin{cases} I(\tau_{ik} = 0) & \text{if } \tau_i^* \leq 0 \\ I(\tau_{ik} = m) & \text{if } m-1 < \tau_i^* \leq m, \;\; m > 0 \end{cases} \tag{8}$$

The conditional density function $f(Y_i, X_i, X_i^*|\tau_i, \beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon)$ can be derived as follows. The changepoints partition the time series for the ith firm into $K+1$ segments, and each segment contributes a factor $f_k(Y_{ik}, X_{ik}, X_{ik}^*|\tau_{ik}, \beta_{ik}, \alpha_i, \gamma, \delta, \Sigma_\epsilon)$ to the likelihood as follows

$$\begin{aligned} f(Y_i, X_i, X_i^*|\tau_i) &= \prod_{k=0}^{K} f_k(Y_{ik}, X_{ik}, X_{ik}^*|\tau_{ik}, \beta_{ik}, \alpha_i, \gamma, \delta, \Sigma_\epsilon) \\ &= \prod_{k=0}^{K} \prod_{t=T_{ik}}^{T_{i(k+1)}-1} \Big[(I(X_{it} = 0, X_{it}^* < 0) + I(X_{it} > 0, X_{it}^* = X_{it})) \\ &\qquad \phi(Y_{it} - X_{it}\beta_{2i} - U_{it}^T \alpha_i - W_{it}^T \gamma - \frac{\sigma_{12}}{\sigma_{22}} \epsilon_{it}^X | 0, \sigma_{1|2}^2) \phi(\epsilon_{it}^X = X_{it}^* - \delta Z_{it} | 0, \sigma_{22}) \Big], \end{aligned}$$

where $I(\cdot)$ is the indicator function; the additive term $(I(X_{it} = 0, X_{it}^* < 0) + I(X_{it} > 0, X_{it}^* = X_{it}))$ enforces the consistency between the observed X_{it} value and the latent X_{it} value; $\phi(\cdot|\mu, \sigma^2)$ stands for the density function for normal distribution with mean μ and variance σ^2.

To help understand the likelihood, it is instructive to consider the simpler case of one changepoint, i.e., $K = 1$ and τ_i is a scalar. The likelihood function under this case can have three forms as follows.

(1) $\tau_{i1} = 0$

This corresponds to the case of immediate stable long-term response from firm i. In our derivation, we write the joint distribution of $f(\epsilon_{it}^Y, \epsilon_{it}^X) = f(\epsilon_{it}^X) f(\epsilon_{it}^Y|\epsilon_{it}^X)$ as in Rossi et al. (2005). In this decomposition, $\epsilon_{it}^Y|\epsilon_{it}^X \sim N(\frac{\sigma_{12}}{\sigma_{22}}\epsilon_{it}^X, \sigma_{1|2}^2)$, where $\sigma_{1|2}^2 = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}$. We then have the density function as follows:

$$\begin{aligned} f(Y_i, X_i, X_i^*|\tau_i = 0) &= \prod_t \Big[(I(X_{it} = 0, X_{it}^* < 0) + I(X_{it} > 0, X_{it}^* = X_{it})) \\ &\qquad \phi(Y_{it} - X_{it}\beta_{1i} - U_{it}^T \alpha_i - W_{it}^T \gamma - \frac{\sigma_{12}}{\sigma_{22}} \epsilon_{it}^X | 0, \sigma_{1|2}^2) \phi(\epsilon_{it}^X = X_{it}^* - \delta Z_{it} | 0, \sigma_{22}) \Big], \end{aligned}$$

In this case, the likelihood from firm i contributes information about the stable long-term effect but no information about the short-term effect.

(2) $0 < \tau_{i1} < T - T_{ie} + 1$

In this case, we have the density function as follows:

$$\begin{aligned} f(Y_i, X_i, X_i^* | \tau_i) = & \prod_{t < T_{ie} + \tau_i} \Big[(I(X_{it} = 0, X_{it}^* < 0) + I(X_{it} > 0, X_{it}^* = X_{it})) \\ & \phi(Y_{it} - X_{it}\beta_{0i} - U_{it}^T \alpha_i - W_{it}^T \gamma - \frac{\sigma_{12}}{\sigma_{22}} \epsilon_{it}^X | 0, \sigma_{1|2}^2) \phi(\epsilon_{it}^X = X_{it}^* - \delta Z_{it} | 0, \sigma_{22}) \Big] \\ & \times \prod_{t \geq T_{ie} + \tau_i} \Big[(I(X_{it} = 0, X_{it}^* < 0) + I(X_{it} > 0, X_{it}^* = X_{it})) \\ & \phi(Y_{it} - X_{it}\beta_{1i} - U_{it}^T \alpha_i - W_{it}^T \gamma - \frac{\sigma_{12}}{\sigma_{22}} \epsilon_{it}^X | 0, \sigma_{1|2}^2) \phi(\epsilon_{it}^X = X_{it}^* - \delta Z_{it} | 0, \sigma_{22}) \Big]. \end{aligned}$$

In this case, the likelihood from firm i contributes information for both the stable long-term and temporary short-term effect.

(3) $\tau_{i1} \geq T - T_{ie} + 1$

In this case, we have the density function as follows:

$$\begin{aligned} f(Y_i, X_i, X_i^* | \tau_i) = & \prod_t \Big[(I(X_{it} = 0, X_{it}^* < 0) + I(X_{it} > 0, X_{it}^* = X_{it})) \\ & \phi(Y_{it} - X_{it}\beta_{0i} - U_{it}^T \alpha_i - W_{it}^T \gamma - \frac{\sigma_{12}}{\sigma_{22}} \epsilon_{it}^X | 0, \sigma_{1|2}^2) \phi(\epsilon_{it}^X = X_{it}^* - \delta Z_{it} | 0, \sigma_{22}) \Big]. \end{aligned}$$

In this case, the likelihood from firm i contributes information about the short-term effect but no information about the long-term effect. Given each possible value of τ_i, we can construct the above likelihood. These unit-level likelihood is then combined with the hierarchical prior distribution of unit-level latent data to draw inference on the likely position of changepoints. Note that according to the above derivation, the likelihood from firms with no counterfeit entry would contribute no information about the short-term and long-term effects except serving as a control group and affecting the estimation of other model parameters. Also note that both cases (1) and (3) are unlikely in our dataset because (1) it usually takes time for a firm to design responding strategy and (2) our panel spans twelve years and all firms, if infringed, have at least five years of observations after counterfeit entry, which is long enough for firms to respond. However, for model completeness, we include these two cases in our model development.

2.2. Between-firm Model

In this subsection, we consider modeling the firm-level parameters and latent variables, $(\alpha_i, \beta_i, \tau_i^*)$, as a function of firm-level characteristics. Numerous studies (e.g., Chintagunta, Jain and Vilcassim

1991, Allenby and Lenk 1994, Allenby and Rossi 1999, Ansari, Jedidi and Jagpal 2000, Bradlow and Rao 2000) have demonstrated the usefulness of modeling heterogeneity. The purpose here is to study factors explaining the interfirm difference on the multi-stage effects and the firms' response times to counterfeit entry. We employ the following multivariate normal heterogeneity model.

$$\begin{bmatrix} \alpha_i \\ \beta_{i0} \\ \cdots \\ \beta_{iK} \\ \tau_{i1}^* \\ \cdots \\ \tau_{iK}^* \end{bmatrix} = \Pi Z_i + \begin{bmatrix} e_i^{\alpha} \\ e_i^{\beta_0} \\ \cdots \\ e_i^{\beta_K} \\ e_i^{\tau_1} \\ \cdots \\ e_i^{\tau_K} \end{bmatrix}, \tag{9}$$

where Π is a $n_r \times n_z$ matrix containing the hyperparameters governing the population distribution of firm-level latent variables, n_r is the number of these latent variables, n_z is the number of variables for firm characteristics (plus an intercept term), and $e_i^B = (e_i^{\alpha}, e_i^{\beta_0}, \cdots, e_i^{\beta_K}, e_i^{\tau_1}, \cdots, e_i^{\tau_K})'$ contains random residuals that are assumed to be jointly multivariate normal as

$$(e_i^{\alpha}, e_i^{\beta_0}, \cdots, e_i^{\beta_K}, e_i^{\tau_1}, \cdots, e_i^{\tau_K}) \sim \text{MVN}\left[(0, 0, \cdots, 0, 0, \cdots, 0), \Sigma_e\right]. \tag{10}$$

As in a typical HB model, the above between-firm model allows one to study the determinants of interfirm difference and to leverage strength from different firms in the estimation of firm-level models. On the other hand, this between-firm model has two important benefits, as compared with the conventional HB model. The first benefit is allowance for regime-switching moderating effects: unlike the traditional hierarchical model, the effect of Z are allowed to vary from regime to regime (e.g., a moderator in one regime but not in another). The second benefit is to provide a framework for us to investigate how fast firms respond to market change and study the determinants of firms' heterogeneous response times. In this approach, the latent response times (τ_i^*) are treated the same as the other firm-level latent characteristics, (α_i, β_i). Therefore the model extends the traditional hierarchical model to incorporate the latent response times as additional dimensions of units' response behaviors. Given the staggering patterns of defensive responses by incumbents as identified in prior literature (Robinson, 1988; Bowman and Gatignon, 1995), the following questions naturally arise: Why do some incumbents choose to respond immediately to competitive entry while others delay their responses? The response time likely represents the heterogeneity of the units, such as firms, in their ability to adapt to the changing environment and is an important dimension of firms' response behaviors to study (Robinson 1988, Smith et al. 1989, Heil and Robinson 1991, Bowman and Gatignon 1995). These previous studies typically use survey data to investigate the questions. As noted by the authors, one methodological limitation of the survey approach relates to various potential response biases in the dependent variable. Our proposed method provides an alternative method that uses field data to infer latent response times from the trajectory of the market response observed over time.

The above Equations (1),(3), (4), (5), (6), (9) and (10) specify a probability model describing the data-generating process. The entire parameter vector is $(\Pi, \gamma, \delta, \Sigma_\epsilon, \Sigma_e)$. The likelihood for these parameters is as follows:

$$L(\Pi, \gamma, \delta, \Sigma_\epsilon, \Sigma_e; X, Y) \propto \prod_i \int \cdots \int f(Y_i, X_i | \tau_i^*, \beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon) f(\alpha_i, \beta_i, \tau_i^* | \Pi, \Sigma_e) d\alpha_i d\beta_i d\tau_i^*,$$

where $i = 1, ..., N$, $f(Y_i, X_i | \tau_i^*, \beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon)$ is specified in Equation (7) and $f(\alpha_i, \beta_i, \tau_i^* | \Pi, \Sigma_e)$ is the density function of Equation (9) and (10).

3. Inference

As shown above, the likelihood for the RC-SEM involves integration and summation over the latent variables X_{it}^*, $\alpha_i, \beta_i, \tau_i^*$ and τ_i respectively, which renders inference based on the direct Maximum Likelihood Estimation or Least Square method intractable. Bai and Perron (1998), Fader et al. (2004), and Perron and Yamamoto (2011) consider modeling and testing of multiple changepoints in the frequentest framework. Extending these frequentest procedures to our application encounters several difficulties. Unlike their applications, the likelihood in our application involves high-dimensional integrations that are hard to evaluate numerically. Further complicating the issue is that with multiple changepoints the number of possible partitions within each panel, and thus the number of terms to evaluate in the likelihood functions increases exponentially. For this reason, Fader et al. (2004) need to restrict analysis to a limited number of changepoints. In contrast, our Bayesian approach avoids evaluating the model likelihood. Specifically, we use the data augmentation technique that augments the parameter vector by the latent data, and then sample from the joint posterior distribution of model parameters and latent variables. This approach requires neither evaluating high-dimensional integrals numerically nor the likelihood with a exploded number of terms. As a result, the Bayesian approach is capable of reducing the computational workload from an exponential rate to a linear rate. Moreover, it is straightforward to make inferences on both the population parameters and latent variables under the Bayesian framework. For example, their estimates and the standard errors can be readily obtained from the posterior draws.

To complete our model, we need to specify the priors for the parameters in the model. Let $\Theta = vec(\Pi')$. We assign priors for the model parameters as follows:

$$\Theta \sim N(\mu_\Pi, \Lambda_\Pi^{-1}), \quad \gamma \sim N(\mu_\gamma, A_\gamma^{-2}), \quad \delta \sim N(\mu_\delta, A_\delta^{-1}), \quad \Sigma_\epsilon \sim IW(\nu_\epsilon, S_\epsilon), \quad \Sigma_e \sim IW(\nu_e, S_e), \quad (11)$$

where $IW(\nu, S)$ stands for an inverse-Wishart distribution with ν degrees of freedom and the scale matrix S. The above distributional forms are chosen for priors because these are conjugate priors for deriving the conditionals in our Gibbs sampler. In our analysis, the constants in the priors are chosen in a way so that these priors are relatively diffuse. The assignment of values for the constants is described in Web Appendix A. With the above specified priors, model specification, and observed data X and Y, the posterior distribution of the parameters and latent data is as follows, up to a constant:

$$\pi(\Pi,\gamma,\delta,\Sigma_\epsilon,\Sigma_e,X^*,\alpha_i,\beta_i,\tau,\tau^*|Y,X) \propto \prod_{i=1}^{N} f(Y_i,X_i,X_i^*|\tau_i,\alpha_i,\beta_i,\gamma,\delta,\Sigma_\epsilon) f(\tau_i|\tau_i^*) f(\tau_i^*,\alpha_i,\beta_i|\Pi,\Sigma_e)$$
$$\cdot\, \pi(\Pi|\mu_\Pi,\Lambda_\Pi^{-1})\pi(\gamma|\mu_\gamma,A_\gamma)\pi(\delta|\mu_\delta,A_\delta)\pi(\Sigma_\epsilon|\nu_\epsilon,S_\epsilon)\pi(\Sigma_e|\nu_e,S_e). \quad (12)$$

Because the analytical expression of the posterior distribution is unavailable, we use MCMC sampling method for model inference with details described in the Web Appendix A.

The above estimation algorithm conditions on a known number of changepoints. In practice, an important issue is to determine the suitable value of K, the number of changepoints. An overly large number of changepoints can lead to imprecise model estimation and degraded inferential performance because unnecessary changepoints can lead to few observations per stage and too much variations in the estimation. On the other hand, an insufficient number of changepoints will not capture the multi-stage effects adequately. It is thus likely to have an optimal value of K. The selection of the number of changepoints can be cast as a model selection problem. In Bayesian framework a well established approach to model selection is Bayes factor (BF). The BF compares two competing models via the ratio of the marginal likelihood under the two models and is applicable for comparing non-nested models, as occurred in the selection of the number of changepoints. When computing BF in our application, we employ the approach of Raftery et al. (2007) that improves and stabilizes the harmonic mean estimator of Newton and Raftery (1994).

4. A Simulation Study

In this section, we conduct a set of simulation experiments to evaluate the performance of different models in repeated samples. Because of space limitation we move the details of simulation study to the Web Appendix B and summarize the main result here. As shown in Table 6, the simulation study shows that the estimation algorithm under the RC-SEM model recovers the true values of the temporary short-term β_1 and stable long-term entry effect β_2 reasonably well. Its RMSEs are smallest among four models across different strength of endogeneity (i.e., the range of values for ρ). In addition, the coverage rates of the credible intervals are closest to the nominal 95% rate, among all methods. The simulation study shows that both the SEM model and OLS that ignores the heterogeneity in latent response times attenuate entry effects. The attenuation bias could be as large as 50% reduction in the true effect size. This shows that in the presence of heterogeneous response time, ignoring the heterogeneity and specifying a common response time can lead to severely biased effect estimates. Moreover, the Bayesian estimator from the RC-SEM model has less variability (i.e. smaller standard error) than that from the SEM, because the RC-SEM model provides better model-fitting by taking into account the latent response times. The RC estimates are biased because of the endogeneity issue. The OLS estimate has serious bias, particulary when the endogeneity is strong, and the 95% credible intervals hardly contain the true effect value.

5. Empirical Analysis

5.1. Data

The dataset is a large national sample that includes 31 branded shoe manufacturers and their counterfeiters in China. Both multinational brands in China and Chinese-originated brands are sampled through the stratified random sampling method (Qian 2008). Twenty-two out of the total 23 large branded firms in China are captured, together with a random sample of smaller ones. Detailed financial statements of each sampled company and their counterfeiters are obtained from a 12-year window from 1993-2004. This is a unique dataset that overcomes severe data limitations common in the underground economics.

Our study uses a natural experiment arising from an emergent diversion of government enforcement resources from fashion products to several other sectors. In the early 1990s, a series of unexpected accidents took place in sectors such as food, drug, gas, etc., due to sub-quality products. The bureau that's in charge of monitoring counterfeits and sampling products, the Quality and Technology Supervision Bureau (QTSB), had to urgently reallocate resources away from monitoring footwear (and other fashion) trademarks to guarantee product safety in these other sectors around the year 1995. Table 1 shows the drastic reduction in government resources allocated to shoe sector monitoring after 1995, along with summary statistics of some other variables. This leaves loopholes for counterfeiters to massively enter the footwear industry (Table 1). The branded companies relied on their own 'brand-protection" offices to monitor the market, to report infringers of their own brands to the QTSB, and to track down counterfeits together with the QTSB. Counterfeits, as these "brand-protection" offices and QTSB shared with me, include all the illegal producers that infringed on the brand by illegally claiming/copying the brand on the counterfeits.

The entry and sale of counterfeits is likely endogenous due to unobserved time-varying firm characteristics. Ignoring the endogeneity, when present, will lead to erroneous inference about its causal effect on the authentic firms' pricing. To identify the effects of the counterfeit entry, we adopt an IV strategy as used in Qian (2008). The identification strategy makes use of the above natural experiment in which the exogenous shocks led to the loosening of the Chinese government's monitoring of footwear trademarks, and exploits the interaction between the unexpected enforcement change and the relationship between each branded company and the government, as proxied by the number of days it took each company to pass the required International Standards (ISO) applications. This IV strategy recognizes that branded companies that have better relationships with the government are less affected by the sudden loosening of trademark enforcement, and hence face less threats by counterfeit entry. The identification strategy uses variations in these IVs to tease out the exogenous components of the counterfeit sales which is then used to identify the causal effect of counterfeit entry. More institutional details regarding the IV validity are discussed in Qian (2008). In Equation (4), we use $LOOSE$, $RELATION$ and $LOOSE * RELATION$ as the main instruments, where $LOOSE$ is an indicator variable denoting the loosening of Chinese government enforcement in monitoring the footwear trademarks, $RELATION$ denotes the number of days it took the company to pass the required ISO applications, and $LOOSE * RELATION$ is the interaction between these two variables. Our analysis consists of two main parts: study and test for the dynamic effects of counterfeits entry (Section 5.2), and investigate the drivers for the interfirm difference in their response behavior (Section 5.5). Neither was formally tested or examined in prior research (Qian 2008).

5.2. Results from a Bayesian Analysis using RC-SEM

We first use the RC-SEM to estimate and test for the dynamic effects of counterfeit entry. In our empirical RC-SEM model, the outcome variable Y_{it} is the logarithm of the deflated prices for the ith authentic firm's high-end product at year t. The explanatory variable of main interest, X_{it}, is the quantity of counterfeit products in the market faced by the ith authentic brand at year t, divided by the sale quantity of this authentic firm. This variable represents a normalized measure of the significance of counterfeit threat that quantifies the relative importance of counterfeit entry. The covariate U in Equation (1) includes firm dummies and W includes year dummies. The year dummies capture the effects of common shocks to the market that may vary by time. The firm dummies capture the effects of unobserved firm-level time-constant characteristics.

Authentic firms respond to their counterfeit entry with various time lags. As a result, the change in prices of their products are manifested in the data only after the response time. In the analysis below, we will use RC-SEM to explicitly model the latent heterogeneous response times among the authentic firms. One advantage of the RC-SEM is that it automatically detects the presence and location of firm-specific changepoints in the outcome variable time-series. We first fit the models in which the covariate Z_i in Equation (9) contains only the intercept. Table 2 reports $2\ln(BF)$ that compares models with different values of K, the number of changepoints occurring in the dataset. A general rule is that a value of larger than 5 for $2\ln(BF)$ provides strong evidence against the null model (Raftery 1996). It shows that the model with one or two changepoints are among the best ones, and the model with $K = 1$ is better than $K = 2$. We also examined the parameter estimates for $K = 2$, which gives similar parameter estimates for the short- and long-term estimates. The difference is that the model under $K = 2$ also shows a somewhat smaller positive intermediate effect for a very short time period, which can be considered as a manifest of model overfitting. Overall, in addition to the comparison based on BF, the difference in parameter estimates between $K = 1$ and $K = 2$ is not substantial from this closer examination of model fitting. In this case, a more parsimonious model is preferred. Therefore the analysis below will use the model with $K = 1$.

For comparison purposes, we also fit three nested models of the RC-SEM. The first one is the OLS with a pre-specified response time common to all firms. This model specification recognizes the nonstationary effects of the counterfeit entry, and thus it is more realistic than a static OLS model that assumes static effect of counterfeit sales. Although convenient, this analytic strategy suffers several drawbacks. First, this approach requires researchers to pre-specify a common value of response time, which is typically not easy to do and requires considerable prior knowledge on the underlying economic activities. It would be preferable to have a data-driven approach to automatically select the timepoint that separates out the short- and long-term effects of market change. In our analysis, we pre-specify the common response time to be two years, which is the value closest to the posterior mean of response times estimated from our RC-SEM shown later. Second and more importantly, it is often the case that there exists a significant amount of heterogeneity among different units (e.g. firms) in their response times (Bowman and Gatignon 1995, Fader et al. 2004). However, the simple method ignores timing heterogeneity, which can substantially bias the effect estimation. Third, specifying a common response time provides no basis to study what relates to the response speed. The second one is a random-changepoints (RC) model. This model recognizes the heterogeneous response time but ignores the endogeneity issue. The third one is the simultaneous equation model (SEM) with a pre-specified response time common to all firms. The standard SEM accounts for the endogeneity issue. However, like OLS, SEM also assumes a common response time with a value of two years. All model fittings run the Gibbs sampler for 50,000 iterations and discard the first 10,000 iterations as the burn-in period. The convergence of the Markov chains to the stationary distributions after the burn-in period is confirmed via examining traceplots and the Geweke's diagnostic statistics (Geweke 1992).

The posterior means and standard deviations of the parameter draws from the Gibbs sampler for all models are reported in Table 3. Under the RC-SEM, we are able to detect the presence of response times that are heterogeneous among the authentic firms. Figure 2 (a) and (b) plots the prior and posterior distributions of the latent response time τ_i^*. As shown in the figure, data provide a substantial amount of information so that the posterior distribution of τ_i^* is a much more condensed distribution as compared with its prior distribution. The average response time is estimated to be 1.6 years with a standard deviation of 1.1. This shows that the authentic firms took considerably different amounts of time to design and implement counter measures against their counterfeits. In contrast, neither the OLS nor SEM provides the estimates of the response times since both methods pre-specify them to be a common value of two years.

The results also show that the effect estimates of counterfeit entry are also different for different methods. Because the covariate Z_i in Equation (9) contains only the intercept, the parameter estimates in Π are the population mean effects. As shown in Table 3, RC-SEM shows that there is a negative (-0.34) short-term population mean effect and a positive (1.61) long-term population mean effect of counterfeit entry on the authentic firms' prices. The 95% credible intervals for both effects exclude zero. In comparison, the results from SEM show attenuated effects for both the short-term and long-term effects. In particular, the 95% credible interval for the short-term effect under the SEM includes zero. This shows that ignoring the heterogeneous response times, like what a standard SEM does, can lead to attenuated effect estimates and lose power to detect a change in the marketing response. RC and OLS also yield different estimates of the effect estimates since they do not model the endogeneous entry or heterogeneous response times. [4] Table 3 shows a negative value of the posterior mean of the covariance term σ_{12}. Figure 2 (c) and (d) plot the prior and posterior distributions of the correlation coefficient, $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$, and show that data provide strong evidence for the presence of a negative correlation, as compared with its prior distribution. This implies that there were some unobserved factors that affected the price and counterfeit entry in opposite directions. These factors could be a firm's managerial skills that are positively correlated with product prices and negatively correlated with the counterfeit entry. It is also possible that higher prices are associated with higher quality products that are harder to be imitated or counterfeited. The endogeneity issue, if not accounted for in the modeling, will lead to inconsistent estimate of causal effects.

We also conduct model comparisons using Bayes factors (Raftery et al. 2007), in which the RC-SEM is considered as the full model, and the other three models (SEM, RC and OLS) are considered as various nested models of the RC-SEM. [5] Our calculation shows that the $2\ln(BF)$ for RC-SEM against SEM (null model) is 321.2 which provides overwhelming evidence for the presence of heterogeneous random changepoints among firms. The $2\ln(BF)$ for RC-SEM against RC (null model) is 117.1, which provides very strong evidence for the presence of endogeneity of counterfeit entry. The $2\ln(BF)$ for RC-SEM against OLS (null model) is 387.8, which provides overwhelming evidence for the simultaneous presence of both heterogeneous random changepoints among firms and the endogeneity of counterfeit entry. Figure 3 presents the posterior predictive

[4] In particular, both models significantly underestimate the long-term effect of counterfeit entry. Furthermore, as shown in the next subsection, these two models also mis-assess the effects of several factors on marketing outcomes and are not able to identify their significant effects.

[5] Because the RC and OLS use a single-equation approach, in order to make their marginal likelihood comparable with those of the SEM-type models, we have added the contribution of the likelihood from an independent model for X in the calculation of marginal likelihood for these two models.

model checking results. The plots show the simple approach that assumes a common changepoint of two years captures the nonstationary effects poorly. It attenuates both the short-term (most significantly) and the long-term effects considerably. As a result, there is appreciable discrepancy between the actual values and the predicted values under the traditional SEM. On the other hand, the RC-SEM model performs much better and provides a substantially improved model fitting.

5.3. Robustness Checks

Proactive Firm Reaction. The analysis above studies firms' responses after their own counterfeits entered market. Because the counterfeit entry time (i.e., T_{ie}) varies among firms, it was possible that firms might take proactive pricing actions after noticing counterfeit entry for other firms and before the entry of their own counterfeit. To provide a more complete picture of firms' response behavior, we conducted an analysis using a model expanded from the above RC-SEM. Specifically, we expand the model for the pre-entry period as dictated in Equation (1) to: $Y_{it} = \alpha_i + x'_{it}\beta_{i0} + W'_{it}\gamma + \epsilon^Y_{it}$ if $t < \min_i(T_{ie})$; $Y_{it} = \alpha_i + \delta_i + x'_{it}\beta_{i0} + W'_{it}\gamma + \epsilon^Y_{it}$ if $\min_i(T_{ie}) \leq t < T_{ie}$, where $\min_i(T_{ie})$ denotes the earliest counterfeit entry time observed in the sample; and the model for post-entry period remain the same as before. The added parameter δ_i captures the potential proactive pricing action for firm i in the period between $\min_i(T_{ie})$ and T_{ie}. The prior for δ_i is assigned using the approach described in Web Appendix A. We then fit the data with this expanded model. The estimate of the population mean of δ_i is small and nonsignificant with its posterior mean (SD) as -0.03 (0.05). We also perform Bayesian model comparison that compares this expanded model with the RC-SEM fitted above. The $2ln(BF)$ for the RC-SEM estimated in the section above against this expanded model (in the denominator when computing $2ln(BF)$) is 6.8. The model estimation and test provide no evidence for systematic price changes in this pre-entry period. In the pre-entry period when the counterfeits entered the market for other firms, it may make sense for unaffected firms to maintain a price sufficiently high to signal the high quality of their authentic products. However, when counterfeits entered massively for the incumbent, this will cause significant price competition effect. In the pre-entry period, it is likely the signaling force is dominant, while in the post-entry period the competition force is dominant. Our analysis also shows that the prices among authentic firms are fairly stable over time in pre-entry period, which suggests that the price competition among authentic firms cannot explain the significant price changes after counterfeit entry. [6]

[6] As explained in Section 5.1, our strategy to identify the counterfeit entry effects uses a natural experiment coupled with the instrumental approach. Our approach can be considered as a "limited information" approach that separates the counterfeit entry effects from other effects such as the competition among authentic firms, as compared with a full information approach that models entire complicated system. Because our main purpose is to identify counterfeiting effects, we use this "limited information" identification approach. As discussed in Chintagunta et al. (2006), the estimates from such a limited information approach are consistent and more robust to the model misspecification as compared with a full-information approach, although the estimates may be less efficient than a full-information approach.

Alternative Definition of X. The above analysis defines the key variable X as the share of counterfeit sales, relative to the concurrent authentic sales. The normalization is meant to measure the relative severity of counterfeit entry. To evaluate whether our analysis is robust to alternative definition, we use the un-normalized counterfeit sales in the unit of million pairs as the X variable and refit the RC-SEM. The RC-SEM using this new X variable shows a statistically significant population mean short-term negative effect, -0.06 (0.02) and positive long-term effect, 0.26 (0.03). The 95% Bayesian credible intervals of the population variance parameters for these effects also exclude zero, indicating significant entry effect heterogeneity. In comparison, the conventional IV analysis using constant two-year cut-off shows no dynamic effects and statistically insignificant short-term effect with its population mean estimate of 0.02 (0.04). The analysis demonstrates the robustness of our findings to alternative definition of the key independent variable.

5.4. Summary and Interpretation of Dynamic Effects Results

The analysis above reveals statistically significant dynamics effects of counterfeit entry: authentic price falls first and then increases, providing a more complete picture of firms' response to counterfeit entry. In particular, the RC-SEM finds a highly significant short-term effect of counterfeit entry, whereas in the traditional SEM this effect is not detectable. Our analysis in the next section finds that the pre-entry quality of authentic firm *moderates* the short-term entry effect, which suggests that the short-term effect is due to price competition of counterfeit entry. We also find that quality strongly *mediates* the long-term price increase. Specifically, when we add the logarithm of the concurrent cost variable $Cost_{it}$ to become one variable in U_{it} in our RC-SEM model specification, we find that the long-term effect dropped from 1.61 to 0.41 . Therefore a majority of long-term price increase (>70%) can be explained away by the quality upgrade. Overall, the dynamic effects estimates are consistent with the theoretical prediction that counterfeit entry has both a price competition effect and an effect of stimulating innovation; This precisely reconcile two strands of theoretical predictions on competition effects in the industrial organization.

5.5. Investigating Factors Moderating Firms' Responses

In this subsection, we conduct a finer hierarchical analysis to study factors relating to authentic firms' responses to the counterfeit entry. The firms' response behaviors studied here include both their response magnitude in short term and long term as well as their response speed. Specifically, we expand the RC-SEM model as specified above and include a set of observed firm characteristics to explain the differential responses in the between-firm model as specified in Equation (9). The descriptive statistics of the firm characteristics included in Z are summarized in Table 4.

An issue commonly ignored in hierarchical Bayesian analysis is the potential endogeneity of moderating variables. There are several considerations that indicate this issue is likely not serious in our application. First, it is important to note that all the moderating variables in our application take on the average of pre-entry values. These average pre-entry firm attribute variables in Z are not expected to correlate significantly, if it exists, with the error terms in the first-stage price equation once the firm-level random effects are included in the model. Second, our interest here is to evaluate how different pre-entry characteristics of firms relate to firms' responses behavior after entry. We use the pre-entry values of market share to proxy for brand popularity so that the usual simultaneity issue associated with price and market share should not apply here. [7] Third, we have included a relatively rich set of variables in our between-firm models. Despite these considerations, the potential endogeneity issue cannot be entirely ruled out. Specifically, there could be nontrivial correlation between Z and the error term e^B in the 2nd stage between-firm model, despite the fact that we include a relative rich set of variables in Z. For example, there could be omitted variables or common persistence in time-varying unobservables not included in Z but may be correlated with variables in Z, resulting in regression-error correlation. Without properly accounting for the potential endogeneity issue, the moderating effect estimates can only be interpreted as correlational, instead of causal relationship. The instrumental variable approach is powerful for controlling for the endogeneity issue. A challenge, however, is the difficulty to find suitable instrumental variables for these firm-level characteristic variables. To overcome this challenge, we apply the approach of latent instrumental variables (LIV) to control for the potential endogeneity issue. The LIV approach has been proposed recently by Ebbes et al. (2005) and successfully applied by Zhang, Wedel and Pieters (2009) in marketing applications. For the most general case, the LIV augments the heterogeneity model in Equation (9) with the following models for the firm-level variables:

$$Z_{ij} = \pi_j' L_{ij} + e_{ij}^Z, \quad j = 1, \cdots, n_Z, \tag{13}$$

where i index firms; j indexes firm-level variables; L_{ij} denotes a single latent discrete instrument with $M (M > 1)$ categories and π_j is a vector of length M for category means. The instrumental variable L_{ij} is uncorrelated with the error terms in the system. To allow for the endogeneity of Z, the error vector $e_i^Z = (e_{i1}^Z, \cdots, e_{in_z}^Z)$ is combined with e^B vector in Equation (9) to form a larger error vector which is modeled as a multivariate normal with mean zero and a new variance-covariance matrix Σ_e. When e^Z and e^B are correlated, the corresponding elements of Z are endogenous. In essence, the LIV approach attempts to approximate the situation where an IV has been identified that partitions the variations in Z into two parts: an explained part by $\pi_j' L_{ij}$ that is exogenous and an unexplained part e_{ij}^Z that is correlated with the other error terms in the system. Ebbes et al. (2005) established the nice properties of the LIV procedure. They then applied the LIV approach

[7] We further control for other potential endogeneity, such as omitted variables or common persistence in time-varying unobservables, through the latent IV approach as detailed later.

to re-evaluate the relationship between education and income, and found that LIV performs well in correcting the bias of the OLS procedure. Their studies also demonstrate the robustness of LIV to the misspecification of the error distributions and to the true number of categories of the instruments. Therefore, in our analysis, we set the number of categories $M = 2$, and assume that the latent binary IV $L_{ij} \sim B(p_L^j)$, where $B(\cdot)$ denotes a Bernoulli distribution and $P_L^j = P(L_{ij} = 1)$. We use $Beta(1,1)$ as the prior for P_L^j. The LIV model are then combined with the RC-SEM model to adjust for the potential endogeneity of variables in Z. The details of MCMC algorithm to sample posterior draws from this combined model is given in Web Appendix A.

In Table 5 we report estimates of the hierarchical parameters in Π for all models considered. The LIV model controls for the potential endogeneity of *Cost*, *Popular*, *BrandNo*, *Ads*, firm-level variables that are either important or suspected to have potential endogeneity concern. The covariates in Z are standardized before entering the hierarchical model for ease of interpretation. We will further explain these firm-level variables and the corresponding results in the following paragraphs.

The authentic product quality could moderate the counterfeiting effects in an important way. Our hypothesis is that it will be easier for consumers to detect counterfeits from authentic products when the authentic quality is higher. Since it is harder for counterfeiters to close the quality gap, the short-term shock to the authentic branded company will be less severe. The analyses do show that the authentic quality, as proxied by unit product costs, helps to alleviate the negative impacts of counterfeit entry on prices in the short run (Column 2 in Table 5). However, because the short-term effect estimates are significantly attenuated in traditional SEM, this effect is not significant under SEM.

Another interesting hypothesis is that counterfeits can serve as free advertising for their authentic counterparts, and help to expand overall demand by invigorating consumers who are otherwise uninterested in branded consumption. If this hypothesis is true, one would expect that the long-term effect tends to increase more for those less-known brands as such free-advertising effects would be more significant for these less-known and less-popular brands. The analysis in the column 3 of Table 5 from RC-SEM-LIV shows a negative and statistically significant moderating effect of brand popularity on long-term price increase, supporting the existence of free-advertising effects of counterfeits. [8]

To test whether the degree of diversification moderates the effects of counterfeiting, we collected data on the number of sub-brands and percentage of the sales values for exports each branded company had. There is no significant effect associated with the number of sub-brands a branded company owns, possibly because counterfeiters infringe on all sub-brands (Column 4 in Table 5).

[8] The negative effect of pre-entry brand popularity on long-term price increase remains statistical significant when we add the logarithm of the concurrent cost variable $Cost_{it}$ to become one variable in U_{it} in our the hierarchical analysis, indicating this effect is not mediated through quality upgrades.

However, companies with larger percentage of sales value for export are less affected by the entry of counterfeits, because they are more diversified than the companies that rely primarily on the domestic market where counterfeits massively entered. They correspondingly have less urgency to respond to counterfeiting and have a longer response time (Column 5 in Table 5).

In theory, the more innovative a company is, the faster it will come up with newer product designs and innovations to differentiate from counterfeits. We use two alternative proxies for innovativeness: annual R&D expenditures and patent application costs of each branded company. R&D expenditures measure more of the inputs to innovation while patent costs proxy for innovation outputs (Qian 2007). These two variables are highly correlated (correlation coefficient = 0.97), so we include only the patent costs in the main specifications. Robustness checks using R&D instead of patent costs yield similar results. As expected, the companies with higher levels of patent costs or R&D responded to counterfeiting in a shorter time-frame by introducing a higher priced high-end shoes, as compared to companies with lower levels of innovativeness (Column 6 in Table 5). This suggests that the innovative companies not only innovate faster in the face of competition, but also innovate with better products given that Qian (2008) has shown very high correspondence between these shoe prices and their unit product costs as well as characteristics.

We additionally have information on the annual advertising expenditure of each branded company. While heavier advertising could imply the firms' strong intention to familiarize consumers with the branded products, it could also present a larger brand premium for counterfeiters to free ride on. It then becomes an empirical question whether advertising moderates the effect of entry by counterfeiters. Column 7 of Table 5 under RC-SEM-LIV did not show any significant effects of pre-entry average advertising expenditure, after controlling for its potential endogeneity and other variables in the model.

Finally, we gathered data on human capital within companies to test whether and how this factor moderates the effects of counterfeiting and brand responses. We include the employment and total annual wages of the branded companies to proxy for brand-level human resources. Wage is a proxy for skilled labor commonly used in the economics literature (Huang *et al.* 2012). Columns 8 and 9 in Table 5 demonstrate that the more human resources a branded company has, the shorter time it takes to respond to counterfeits by innovating.

One important thing to note in the analysis is that the traditional SEM and OLS method do not model the heterogeneous response times and therefore do not allow for the study of what affects firms' reaction speed. Furthermore, ignoring the heterogeneous response time also lead to mis-assessment of the effects of firm characteristics and some important moderators are not identified in SEM and OLS, but in RC-SEM and RC-SEM-LIV. The RC model allows for this feature but does not model the endogeneity issue. Its effect estimates are also different from those from the RC-SEM(-LIV).

6. Discussion

With the growing emerging markets and globalization, counterfeiting has gained pervasiveness and presented worldwide impacts. It is therefore necessary to study the impacts of counterfeiting and its policy implications. Our paper aims to address this important substantive issue through our unique dataset and careful analyses. Like many other empirical studies, the issues of *endogeneity*, *discontinuity in response*, *heterogeneity* and *regime-switching* moderating effects are encountered and can potentially spoil inference if not accounted for properly. In this paper, we consider the modeling and inference of a simultaneous equation model with heterogeneous changepoints, and apply the model to study the multi-stage causal effects of counterfeit entry on the authentic firm's price. Our application and simulations demonstrate the importance of accounting for the aforementioned data features. Because the proposed method considers both the heterogeneous nature of marketing players' response times and the endogeneity issue, it minimizes the bias in the effect estimation. Using the RC-SEM method that permits changepoints that are heterogeneous among firms, we are able to test and identify dynamic effects of counterfeit entry. In particular, the new model identifies compelling evidence for (1) the short-term negative effects of counterfeit entry, which was masked in the conventional SEM that ignores the heterogeneity in response timing, and (2) significantly larger long-term positive effects. These improved effect estimates are critical to identify causal drivers that drives the dynamic effects and important factors that explains inter-firm response differences. We find that improvement in the effect estimation provided in RC-SEM leads to substantially more power to identify factors moderating the entry effects, some of which are not identified in traditional HB models. These more accurate short-term and long-term effect estimates are also beneficial for future welfare analyses and policy experiments.

The empirical results in this paper unify two strands of Industrial Organization literature on the entry effects on prices. In particular, the finding that the authentic prices fell immediately upon the entry of counterfeiters can be explained by Fudenberg and Tirole (2000). That is, new entry imposes competitive pressure in the short-run. Our finding that pre-entry quality moderates the short-term counterfeit entry effects supports this theory. We further identify that authentic prices rose substantially on average two years after counterfeit entry. This positive effect could be resolved with the other strand of theories. Notably, Frank and Salkever (1997) predicts that generic entry could steal away the price-sensitive consumer segment, leaving behind a more inelastic demand for the branded companies to re-optimize into a higher price. Qian (2006) predicts that companies invest to differentiate their products from counterfeits through innovation, self-enforcement, vertical integration, as well as price signaling, and all these mechanics lead to price increases. Our hierarchical analysis also reveals that the long-term price increases more for those less-known and less-popular brands, providing evidence for the free-advertising effects of counterfeits. It becomes apparent that these theories can better explain the long-term entry effects, and are complementary rather than contradictory to traditional economic theories that predict negative price shocks. The empirical findings on the pricing effects of counterfeiting can also shed lights on the private label literature, where imitation or copycat strategy accounts for more than 50% of the store brand introductions (Scott Morton and Zettelmeyer, 2004).

The improved hierarchical analyses also uncover a set of firm characteristics that moderate the timing and magnitude of pricing responses to entry by counterfeiters, and prescribe effective brand management strategies tailored to each type of firms. For example, the finding that pre-entry quality reduces the harmful effects of counterfeit entry provides direct evidence for the important role of product quality as a management strategy against counterfeits. The finding of the free advertisement effect of counterfeit entry can also have important management implications for authentic firms to deal with counterfeiting. It suggests that there may be an optimal enforcement level to control for the amount of counterfeits, and the optimal enforcement level will differ by firms' brand popularity levels. Our modeling framework also provides a new approach to study firms' response time, an important dimension of firms' response behavior. We find that firms that have more human resources or less diversification from markets affected by trademark infringement are quicker to fire sustainable long-term responses to new competitive threats by counterfeiters.

The need to account for endogeneity, structural changes and heterogeneity simultaneously is not limited to our dataset. Our modeling framework makes several notable methodological contributions to address these empirical issues. It extends two important business analytic tools, Hierarchical Bayesian models and Random Changepoints models, in several important ways as summarized in the Introduction section. In particular, although the random changepoints model has become an emerging business analytic tool and has been increasingly used for business applications, a stringent assumption is the exogeneity assumption that limits its applicability. [9] Our approach relaxes this important assumption and has the potential to substantially expand the scope of its use.

References

Allenby, G. M. and Lenk, P. (1994) "Modeling Household Purchase Behavior With Logistic Normal Regression," *Journal of the American Statistical Association,* **89**, 1218–1231.

Allenby, G. M., and Rossi, P.E. (1999), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics,* **89**, 57–78.

Almon, S. (1965), "The Distributed Lag Between Capital Appropriations and Expenditures," *Econometrica,* **33**, 178–196.

[9] For example, as discussed in Fader et al. (2004), both the distribution build and competitive effects could lead to endogeneity issues in new product sales forecasting.

Ansari, A., K. Jedidi and S. Jagpal (2000) "A Hierarchical Bayesian Approach for Modeling Heterogeneity in Structural Equation Models," *Marketing Science*, 328–347

Bai, J (1997), "Estimation of a Change Point in Multiple Regression Models," *Review of Economics and Statistics,* **79**, 551–563.

Bai, J. and Perron, P. (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, **66**, 47-78.

Böckenholt, U. and Dillon W.R. (1997). "Some New Methods for an Old Problem: Modeling Preference Changes and Competitive Market Structures in Pretest Market Data." *Journal of Marketing Research,* **15**, 130-142.

Bowman, D., H. Gatignon (1995). "Determinants of competitor response time to a new product introduction." *Journal of Marketing Research.* **32** (February): 4253.

Bradlow, E.T. and Rao, V.R. (2000), "A Hierarchical Bayes Model for Assortment Choice," *Journal of Marketing Research*, Vol. XXXVII (May), 259-268.

Carlin, B., Gelfand, A.E. and Smith, A.F.(1992), "Hierarchical Bayesian Analysis of Changepoint Problems," *Applied Statistics,* **41**, 389–405.

Carlton D. and Perloff J. M. (2005), *Modern Industrial Organization*, 2nd Edition, Pearson/Addison Wesley.

Chib, S. (1998) "Estimation and comparison of multiple change-point models," *Journal of Econometrics,* **86**, 221–241

Chintagunta, P., Jain, D., and Vilcassim, N. (1991)"Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data," *Journal of Marketing Research*, 417–28.

Chintagunta, P. K., Erdem, T., Rossi, P. E., and Wedel, M. (2006) "Structural modeling in marketing: Review and assessment," *Marketing Science*, **25**, 604-616.

Choate, P. (2005). *Hot Property : The Stealing of Ideas in an Age of Globalization.* Westminster, MD, USA: Alfred A. Knopf Incorporated, 2005.

Deleersnyder, Barbara, Marnik G. Dekimpe, Jan-Benedict Steenkamp and Oliver Koll (2007), "Win-win Strategies at Discount Stores", *Journal of Retailing and Consumer Services*, **14**(5), 309-18.

Ebbes, P., Wedel, M., Böckenholt, U., and Steerneman, T. (2005), "Solving and Testing for Regressor-Error (in)Dependence When no Instrumental Variables are Available: With New Evidence for the Effect of Education on Income,", *Quantitative Marketing and Economics,* **3**, 365–392.

Fader, P.S., Hardie, B.G.S. and Huang, C.Y. (2004), "A Dynamic Changepoint Model for New Product Sales Forecasting,", *Marketing Science,* **23**, 50–65.

Fong, D.K.H. and DeSarbo, W.S. (2007), "A Bayesian Methodology for Simultaneously Detecting and Estimating Regime Change Points and Variable Selection in Multiple Regression Models for Marketing Research,", *Quantitative Marketing and Economics,* **5**, 427–453.

Frank, R.G. and Salkever, D.S. (1997), "Generic Entry and the Pricing of Pharmaceuticals", *Journal of Economics & Management Strategy,* **6:1**, 75–90.

Fudenberg, D. and Tirole, J. (2000), "Pricing a Network Good to Deter Entry", *The Journal of Industrial Economics,* **48:4**, 373–390.

Heil, O., and Robertson, T.S. (1991), "Toward a Theory of Competitive Market Signaling: A Research Agenda," *Strategic Management Journal,* **12**, 403–418.

Hoch, Stephen and Shumeet Banerji (1993), "When Do Private Labels Succeed?" *Sloan Management Review***34**(4), 57-67.

Huang, Yasheng, Li Jin, and Yi Qian (2012), "Does Ethnicity Pay?" *Review of Economics and Statistics*, forthcoming.

Kotler, P. (1971), "Marketing Decision Making: A Model Building Approach," Holt, Rinehart & Winston.

Lachaab, M., Ansari, A., Jedidi,K., and Trabelsi, A. (2006), "Modeling Preference Evolution in Discrete Choice Models: A Bayesian State-Space Approach," *Quantitative Marketing and Economics*,**4**, 57–81.

Neelamegham, R., and Chintagunta, P.K. (2004), "Modeling and Forecasting the Sales of Technology Products," *Quantitative Marketing and Economics*,**2**, 195–232.

Netzer, O., Lattin, J. M., and Srinivasan, V. (2008), "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science*,**27**, 185–204.

Newton, M. and Raftery, A. (1994) "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society, Series B,* **56**, 3-48.

Pauwels, Koen and Shuba Srinivasan (2004), "Who Benefits From Store Brand Entry?", *Marketing Science*, **23**(4), 596-610.

Perron, P., and Yamamoto, Y. (2011), "Estimating and Testing Multiple Structural Changes in Models with Endogenous Regressors," *Working Paper*, Boston University.

Poulsen, C.S. (1990), "Mixed Markov and Latent Markov Modeling Applied to Brand Choice Behavior," *International Journal of Research in Marketing*,**7**, 5–19.

Qian, Y. (2006), "Pricing and Marketing Effects of Entry by Imitators or Counterfeits", *Harvard WP*.

Qian, Y. (2007), "Do National Patent Laws Stimulate Domestic Innovation in a Global Patenting Environment? A Cross-Country Analysis of Pharmaceutical Patent Protection, 1978-2002," *The Review of Economics and Statistics,* **89(3)**, 436–453

Qian, Y. (2008), "Impacts of Entry by Counterfeiters," *Quarterly Journal of Economics*, **123(4)**, 1577–1609.

Raftery, A., Newton, M., Satagopan, J.M., and Krivitsky, P.N. (2007), "Estimating the Integrated Likelihood via Posterior Simulation Using the Harmonic Mean Identity, " *Bayesian Statistics,* **8**, 1–45.

Ramaswamy, V. (1997). "Evolutionary Preference Segmentation with Panel Survey Data: An Application to New Products," *International Journal of Research in Marketing*. **14**: 57–80.

Rao, Akshay R. and Kent B. Monroe (1996), “Causes and Consequences of Price Premiums”, *Journal of Business*, **69**, 511-35.

Robinson, W. T. (1988). “Marketing Mix Reactions to Entry.” *Marketing Science.* **7 (4)**: 368–385.

Schweidel, D.A. and Fader, P.S. (2009), “Dynamic Changepoints Revisited: An Evolving Process Model of New Product Sales ,” *International Journal of Research in Marketing,* **26**, 119-124.

Scott Morton, Fiona and Florian Zettelmeyer (2004), “The Strategic Positioning of Store Brands in Retailer-manufacturer Negotiations”, *Review of Industrial Organization*, **24**, 161-94.

Smith, K.G., Grimm, C.M., Chen, M., and Gannon, M. J. (1989), “Predictors of Response Time to Competitive Strategic Actions: Preliminary Theory and Evidence,” *Journal of Business Research,* **18**, 245–258.

Van Heerde, H.J., Mela, C.F., and Manchanda, P. (2004), “The Dynamic Effect of Innovation on Market Structure,” *Journal of the Marketing Research,* **XLI**, 166–183.

World Customs Organization (2004), “Counterfeiting Congress Calls for Public-private Cooperation”. http://www.wcoomd.org/ie/En/Press/press.html, May 25, 2004.

Zhang, J., Wedel, M., and Pieters, R. (2009), “Sales Effects of Attention to Feature Advertisements: A Bayesian Mediation Analysis,” *Journal of the Marketing Research,* **XLVI**, 669–681.

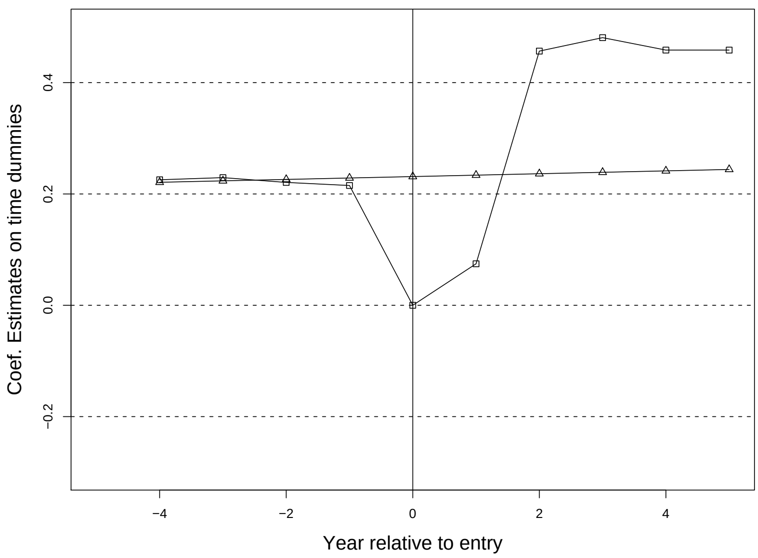

Figure 1 Average log deflated authentic price versus the years relative to the counterfeit entries. −□−: the regression coefficient estimates on a set of time dummies denoting the number of years relative to the counterfeit entries with the log deflated price for high-end product as the response variable. −△−: prediction based on the pre-entry price trend. The year of entry is anchored at the origin.

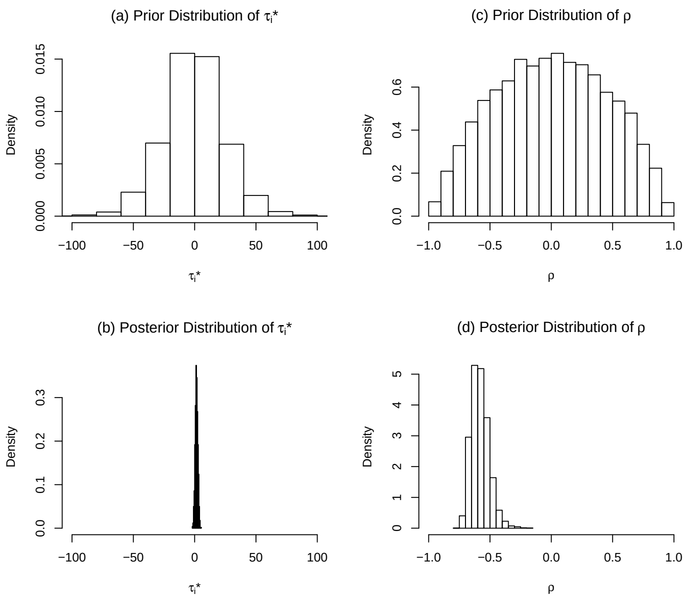

Figure 2 **Comparison of Prior and Posterior Distributions of τ_i^* and ρ.**

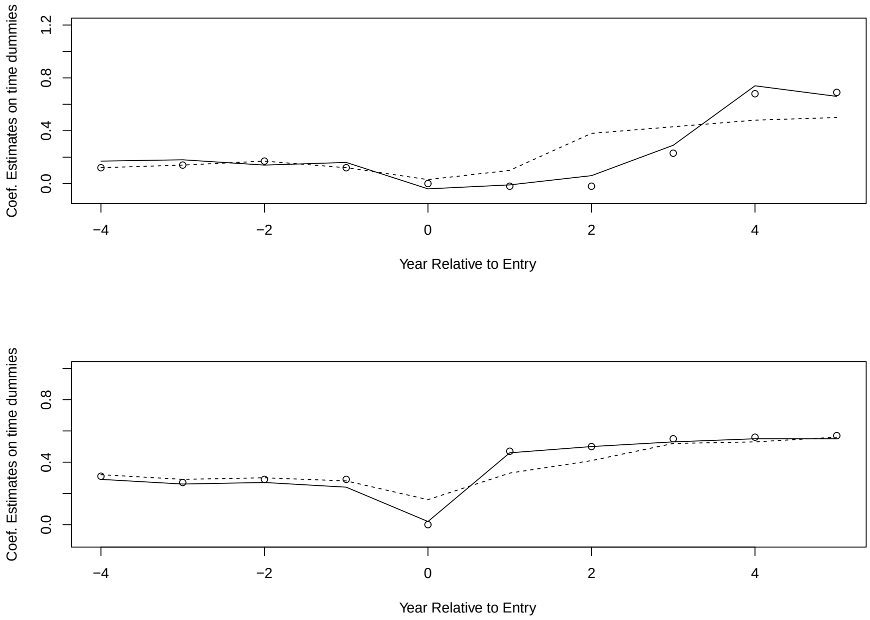

Figure 3 **Posterior Predictive Model Checking. The circles represent the regression coefficient estimates on the set of time dummies denoting the number of years relative to the counterfeit entries with the original data on the log deflated price for high-end product as the response variable. The solid line represents the estimates of the same regression coefficients, obtained as the average over 1000 posterior estimates generated from the RC-SEM. The dotted line is obtained from the SEM assuming a common changepoint of two years. The upper panel is obtained for firms with the posterior mean of τ_i^* larger than two years, and the lower panel for posterior mean τ_i^* smaller than two years. The year of entry is anchored at the origin.**

Table 1 Summary Statistics Pre and Post the Policy Change.

Variable	Pre-1995	Post-1995
Share of Government Resources	0.11	0.02
devoted to shoe sector monitoring	(0.004)	(0.001)
Workdays authentic company took to	142	149
pass ISO (relationship proxy)	(116.5)	(112.6)
Authentic firm Brand-protection Office	0.17	4.00
personnel (head count)	(0.46)	(2.23)
Fake Sale quantity (in 10,000 pairs)	Median 0	85.71
	Range: 1.2-1.9	(75.85)
Authentic Sale Quantity (in 10,000 pairs)	309.38	558.28
	(725.76)	(995.82)
Fake shoe Price (deflated, in US $)	Median 0	7.32
	Range 8.33-10.4	(4.2)
Fake shoe cost (deflated, in US $)	Median 0	2.66
	Range 2.2-3.6	(1.56)
Authentic high-end shoe price	43.3	61.5
	(20.5)	(40.6)
Authentic high-end shoe cost	33.5	47.0
	(19.1)	(30.0)
Authentic medium-end shoe price	27.3	32.3
	(10.2)	(18.7)
Authentic medium-end shoe cost	19.6	24.7
	(9.0)	(13.5)
Authentic low-end shoe price	16.9	18.8
	(8.2)	(10.5)
Authentic low-end she cost	14.8	14.4
	(6.7)	(7.7)
N	62	310

Note: Table presents mean and standard deviation (in parentheses) unless noted otherwise. Prices and costs are deflated using the Consumer Price Index published in the WDI (Year 1995 set as the base year in the database.)

Table 2 Comparison of Models with Different Numbers of Changepoints.

K	$2\ln(BF)$
0	488.3
1	—
2	4.2
3	69.2

Note: All the models are compared to the model with $K=1$, which enters as the numerator in the computation of BF.

Table 3 Estimation Results Under RC-SEM and Different Reduced Models When Z Contains Only Intercept.

Explanatory Variable	RC-SEM	SEM	RC	OLS
(1) Model for $log(dfph)$				
Constant	1.36 (0.12)	1.34 (0.10)	1.41 (0.11)	1.36 (0.11)
fksh_{ST}	-0.34 (0.08)	-0.05 (0.17)	-0.57 (0.10)	-0.24 (0.07)
fksh_{LT}	1.61 (0.25)	1.02 (0.24)	1.24 (0.27)	0.75 (0.11)
ResponseTime	1.57 (0.11)	NA	1.65 (0.13)	NA
(3) Model for $fksq$				
Constant	-0.29 (0.10)	-0.32(0.12)	NA	NA
Loose	0.59 (0.10)	0.53 (0.09)	NA	NA
RELATION	0.001 (0.001)	0.001(0.001)	NA	NA
LOOSE*RELATION	0.003(0.001)	0.003 (0.001)	NA	NA
(4) Correlation Σ_ϵ				
σ_{11}	0.01 (0.001)	0.02 (0.002)	0.01 (0.001)	0.02 (0.002)
σ_{12}	-0.011 (0.002)	-0.01(0.002)	NA	NA
σ_{22}	0.026 (0.003)	0.025(0.003)	NA	NA
Year Fixed Effects	Y	Y	Y	Y
No. of Obs.	372	372	372	372

Note: The table lists the posterior mean and standard deviation of model parameters. $log(dfph)$: the logarithm of deflated authentic high-end prices. $fksh$: the quantity of counterfeit products in the market faced by the corresponding authentic firm, as a share of the sale quantity of this authentic firm. $fksh_{ST}$ and $fksh_{LT}$ refer to its short-term and long-term effects, respectively. RC-SEM: the simultaneous equation model with random changepoints. SEM: the standard simultaneous equation model with a common response time of two years. RC: random-changepoint model. OLS: the standard OLS model with a common response time of two years. All models use year fixed effects.

Table 4 Definition and Summary Statistics of Pre-entry Variables of the Authentic Firms.

Variable	Definition	Mean	SD
Cost	Unit Product Cost of High-end Product (US $)	33.53	19.14
Popular	Brand Popularity proxied by market share (%)	2.9	3.7
BrandNo	The Number of Sub-brands	1.45	0.85
Export	Percentage of Sale Values for Export (%)	18.1	12.5
PatCost	Patent Application Costs (US $)	2453.6	1560.1
Ads	Annual Advertisement Expenditure (US $)	1,497,700	2,724,200
Employ	The Number of Employees	813.7	482.6
AW	Total Annual Wages (US $)	482.8	272.2

Table 5: Estimation Results Under RC-SEM and Different Reduced Models When Z Contains Pre-Entry Moderating Variables.

Attribute	Constant	Cost	Popular	BrandNo	Export	PatCost	Ads	Employ	AW
					(1) RC-SEM-LIV				
Constant (α_i)	1.39**	0.43**	0.12	0.01	-0.07	0.08	-0.08	-0.10**	-0.10**
	(0.05)	(0.06)	(0.08)	(0.12)	(0.06)	(0.07)	(0.19)	(0.05)	(0.05)
fksh$_{ST}$ (β_{0i})	-0.29**	0.23**	0.25	-0.10	0.05	0.05	0.16	0.14*	0.24
	(0.12)	(0.09)	(0.16)	(0.13)	(0.06)	(0.11)	(0.32)	(0.09)	(0.14)
fksh$_{LT}$ (β_{1i})	1.76**	0.34	-0.67**	0.69	-0.21	0.17	0.32	-0.32	0.15
	(0.27)	(0.26)	(0.34)	(0.40)	(0.17)	(0.32)	(0.46)	(0.22)	(0.39)
ResponseTime (τ_i^*)	1.41**	-0.10	-0.05	0.03	0.37**	-0.25*	-0.17	-0.35**	-0.27
	(0.17)	(0.12)	(0.22)	(0.34)	(0.11)	(0.14)	(0.26)	(0.13)	(0.21)
					(2) RC-SEM				
Constant (α_i)	1.40*	0.52*	0.07	0.01	-0.05	0.06	-0.04	-0.13*	-0.09*
	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.04)
fksh$_{ST}$ (β_{0i})	-0.25*	0.14*	0.17	-0.01	0.02	0.02	-0.13	0.07	0.15
	(0.08)	(0.06)	(0.10)	(0.04)	(0.07)	(0.08)	(0.15)	(0.08)	(0.12)
fksh$_{LT}$ (β_{1i})	1.90*	0.21	-0.80*	0.23	-0.27	0.23	0.75*	-0.36*	0.29
	(0.21)	(0.18)	(0.25)	(0.17)	(0.15)	(0.23)	(0.36)	(0.17)	(0.31)
ResponseTime (τ_i^*)	1.51*	-0.03	-0.01	-0.15	0.41*	-0.19	-0.06	-0.32*	-0.26
	(0.14)	(0.12)	(0.17)	(0.12)	(0.12)	(0.11)	(0.17)	(0.13)	(0.20)
					(3) SEM				
Constant (α_i)	1.33*	0.54*	0.07	-0.00	0.02	0.06	0.06	-0.14*	-0.10*
	(0.05)	(0.05)	(0.07)	(0.04)	(0.05)	(0.07)	(0.08)	(0.05)	(0.05)
fksh$_{ST}$ (β_{0i})	0.08	0.06	0.01	0.03	-0.02	0.07	0.19	0.09	-0.15
	(0.10)	(0.06)	(0.10)	(0.06)	(0.07)	(0.08)	(0.17)	(0.07)	(0.12)
fksh$_{LT}$ (β_{1i})	1.16*	0.05	-0.09	0.19*	-0.08	-0.15	0.36	-0.07	0.16

continued on next page

Table 5: *continued*

Attribute	Constant	Cost	Popular	BrandNo	Export	PatCost	Ads	Employ	AW
	(0.15)	(0.11)	(0.16)	(0.09)	(0.11)	(0.15)	(0.28)	(0.12)	(0.21)
ResponseTime (τ_i^*)	NA	NA	NA	NA	NA	NA	NA	NA	NA
	NA	NA	NA	NA	NA	NA	NA	NA	NA
(4) RC									
Constant (α_i)	1.45*	0.54*	0.11	0.02	-0.05	0.03	-0.06	-0.13*	-0.11*
	(0.05)	(0.05)	(0.07)	(0.04)	(0.05)	(0.07)	(0.07)	(0.05)	(0.05)
fksh$_{ST}$ (β_{0i})	-0.61*	0.10	0.21*	-0.04	-0.02	0.06	-0.18	0.13	0.16
	(0.11)	(0.07)	(0.10)	(0.07)	(0.07)	(0.09)	(0.14)	(0.08)	(0.11)
fksh$_{LT}$ (β_{1i})	1.24*	0.12	-0.48*	0.18	-0.16	0.23	0.76*	-0.23	0.19
	(0.20)	(0.15)	(0.21)	(0.14)	(0.15)	(0.19)	(0.27)	(0.17)	(0.27)
ResponseTime (τ_i^*)	1.52*	-0.12	0.03	-0.15	0.25*	-0.15	-0.08	-0.25*	-0.15
	(0.10)	(0.10)	(0.12)	(0.09)	(0.09)	(0.13)	(0.15)	(0.11)	(0.18)
(5) OLS									
Constant (α_i)	1.41*	0.54*	0.08	0.01	0.02	0.05	-0.02	-0.14*	-0.11*
	(0.05)	(0.05)	(0.07)	(0.04)	(0.05)	(0.07)	(0.08)	(0.05)	(0.05)
fksh$_{ST}$ (β_{0i})	-0.21*	0.01	0.02	0.04	-0.13	0.04	0.16	0.06	-0.12
	(0.09)	(0.07)	(0.10)	(0.06)	(0.07)	(0.08)	(0.17)	(0.07)	(0.12)
fksh$_{LT}$ (β_{1i})	0.96*	0.04	-0.15	0.21*	0.02	0.04	0.44	-0.07	0.15
	(0.15)	(0.11)	(0.16)	(0.09)	(0.11)	(0.15)	(0.31)	(0.13)	(0.20)
ResponseTime (τ_i^*)	NA	NA	NA	NA	NA	NA	NA	NA	NA
	NA	NA	NA	NA	NA	NA	NA	NA	NA

Note: The table lists the posterior means and standard deviations of model parameters. $log(dfph)$: the logarithm of deflated authentic high-end prices in Chinese Yuan, using the Consumer Price index published in the World Bank World Development Indicators (WDI) (Year 1995 was set as the base year in the database, i.e. CPI=100 in 1995). $fksh$: the quantity of counterfeit products in the market faced by the corresponding authentic firm, divided by the sale quantity of this authentic firm. $fksh_{ST}$ and $fksh_{LT}$ refer to its short-term and long-term effects, respectively. RC-SEM-LIV: RC-SEM with the LIV model adjusting for the potential endogeneity of moderating variables. RC-SEM: the simultaneous equation model with random changepoints. SEM: the standard simultaneous equation model assuming a common response time of two years. RC: random-changepoints model. OLS: the standard OLS model assuming a common response time of two years. "**" and "*" indicate that 95% or 90% credible intervals excludes zero, respectively.

Appendix. Web Appendix

Web Appendix A: Prior Specification and MCMC algorithm for the RC-SEM

In this Online Appendix, we present the details of prior specification and the MCMC algorithm for estimating the proposed simultaneous equation model with random changepoints (RC-SEM).

Equation (11) presents the forms of the priors. In our analysis, we set the constants in the priors as follows: $\mu_\Pi, \mu_\gamma, \mu_\delta$ are assigned as vectors of zeros. $\Lambda_\Pi = 0.01 \times I_{n_r \times n_z}$, $A_\gamma = 0.01 \times I_{n_\gamma}$, $A_\delta = 0.01 \times I_{n_\delta}$, $\nu_\epsilon = 5$ and $\nu_e = n_r + 3$, $S_\epsilon = \nu_\epsilon \Sigma_{\epsilon 0}$ and $S_e = \nu_e \Sigma_{e0}$, where n_r is the dimension of the square matrix of Σ_e, and n_z is the number of columns of Z, n_γ and n_δ is the length of γ and δ. When choosing the value for $\Sigma_{\epsilon 0}$ and Σ_{e0}, we follow the approach suggested in Rossi et al. (2005). Because $\Sigma_{\epsilon 0}$ is related to the mean of the prior for the variance-covariance matrix of the error terms of the simultaneous equation model, we would like to take into account the scale of the outcomes and the explanatory power of the regressors in the assignment of its value. Rather than assigning an arbitrary value, such as an identity matrix, we set $\Sigma_{\epsilon 0} = \hat{\Sigma}_\epsilon^{OLS}$, where the diagonal entries of $\hat{\Sigma}_\epsilon^{OLS}$ are OLS estimates of error variances for Y equation and X equation, separately, and the off-diagonal entries of $\hat{\Sigma}_\epsilon^{OLS}$ are zeros. The choice of the prior is reasonable because the resulting prior is reasonably flat over a wide range of plausible values of the correlation coefficient ρ between ϵ^X and ϵ^Y, the measure of the endogeneity strength. Figure 2 (c) displays the marginal prior distribution of the correlation coefficient $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$ given the choices of the above chosen prior. The histogram is constructed from sampling 10000 iid draws from the priors of Σ_ϵ, and then calculating ρ for each draw. The histogram shows that for this prior, the distribution of correlation coefficient ρ is centered at zero and reasonably spread out between -1 and +1. Similarly, we have set the block in Σ_{e0} for (α_i, β_i) as $\hat{\Sigma}_e^{OLS}$. Because the OLS does not model the heterogeneity of τ_i^*, we set the corresponding diagonal entries in Σ_{e0} to be one. Using a larger value (eg., 2) or a smaller value (e.g., 0.5) has little impact on the resulting estimation.

Given the above prior and posterior distribution derived in Equation (12), we devise a Gibbs sampler to obtain draws from the posterior distribution. The full conditionals of model unknowns for each step of the Gibbs sampler are derived below.

1. Draw $\beta_i, \alpha_i | \delta, \gamma, \Sigma_\epsilon, T_i, X^*$

We decompose the joint bivariate normal distribution of the error term $(\epsilon_{it}^X, \epsilon_{it}^Y)$ as the product of the marginal distribution of ϵ_{it}^X and the conditional distribution of $\epsilon_{it}^Y | \epsilon_{it}^X$. We note that the conditional distribution $\epsilon_{it}^Y | \epsilon_{it}^X$ is $N(\frac{\sigma_{12}}{\sigma_{22}}\epsilon_{it}^X, \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}})$. We write $\epsilon_{it}^Y = \frac{\sigma_{12}}{\sigma_{22}}\epsilon_{it}^X + e_{it}$, where $e_{it} \sim N(0, \sigma_e^2)$, $\sigma_e^2 = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}$ and $e_{it} \perp \epsilon_{it}^X$. Given δ, the error term $\epsilon_{it}^X = X_{it}^* - \delta Z_{it}$. To make notation more compact at the unit level, let $Y_i = (Y_{i1}, ..., Y_{iT}), \epsilon_i^X = (\epsilon_{i1}^X, ..., \epsilon_{iT}^X), e_i = (e_{i1}, ..., e_{iT})$. Then at the unit level, we have:

$$Y_i - W_i'\gamma - \frac{\sigma_{12}}{\sigma_{22}}\epsilon_{it}^X = H_i \begin{bmatrix} \beta_{0i} \\ \cdots \\ \beta_{Ki} \\ \alpha_i \end{bmatrix} + e_i, \tag{14}$$

where the design matrix for the ith unit H_i is:

$$H_i = (H_{i1}', ..., H_{iT}')' = \begin{bmatrix} X_{i1} & . & 0 & U_{i1} \\ . & . & . & ... \\ X_{i,T_{i1}-1} & . & 0 & ... \\ . & . & . & ... \\ 0 & . & X_{i,T_K} & ... \\ . & . & . & ... \\ 0 & . & X_{i,T} & U_{iT} \end{bmatrix}, \tag{15}$$

and $T_{ik} = T_{ie} + \sum_{j=1}^{k} \tau_{ij}$. We rewrite the variance-covariance matrix, Σ_e, in the distribution of the latent variables $(\beta_i, \alpha_i, \tau_i^*)$ in Equation (10) as follows,

$$\Sigma_e = \begin{bmatrix} \Sigma_{e,11} & \Sigma_{e,12} \\ \Sigma_{e,12} & \Sigma_{e,22} \end{bmatrix} \tag{16}$$

where $\Sigma_{e,11}$ and $\Sigma_{e,22}$ are variance-covariance matrix of (β_i, α_i) and τ_i^*, respectively, and $\Sigma_{e,12}$ is covariance matrix between these two blocks of parameters. The prior distribution of $(\beta_i, \alpha_i)|\tau_i^*$ is $N(\mu_{1|2,i}, A_{1|2}^{-1})$ where

$$\mu_{1|2,i} = \mu_{1i} + \Sigma_{e,12}\Sigma_{e,22}^{-1}(\tau_i^* - \mu_{2i})$$
$$A_{1|2}^{-1} = \Sigma_{e,11} - \Sigma_{e,12}\Sigma_{e,22}^{-1}\Sigma_{e,21},$$

and μ_{1i} and μ_{2i} are the prior means of (β_i, α_i) and τ_i, respectively, and $(\mu_{1i}, \mu_{2i}) = \Pi Z_i$. Note that the above conditional mean and variance-covariance matrix should also condition on e^Z if LIVs are used for Z using the results derived in Step 7. Then the conditional draws of (β_i, α_i) can be obtained from the following normal distribution:

$$N\left((\frac{H_i' H_i}{\sigma_e^2} + A_{1|2})^{-1}(\frac{H_i' R_i}{\sigma_e^2} + A_{1|2}\mu_{1|2,i}), \quad (\frac{H_i' H_i}{\sigma_e^2} + A_{1|2})^{-1} \right)$$

where $R_i = Y_i - W_i'\gamma - \frac{\sigma_{12}}{\sigma_{22}}\epsilon_i^X$.

2. $\gamma|\delta, \beta_i, \alpha_i, \Sigma_\epsilon, T, X^*$

Given the prior distribution for γ as $N(\mu_\gamma, A_\gamma^{-1})$, we have the conditional draw of γ as

$$N\left((\frac{W_i' W_i}{\sigma_e^2} + A_\gamma)^{-1}(\frac{W_i' R_i^\gamma}{\sigma_e^2} + A_\gamma\mu_\gamma), \quad (\frac{W_i' W_i}{\sigma_e^2} + A_\gamma)^{-1} \right),$$

where $R_i^\gamma = Y_i - H_i[\beta_i, \alpha_i]' - \frac{\sigma_{12}}{\sigma_{22}}\epsilon_i^X$.

3. $X_{it}^*|\beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon, T$

If $X_{it} > 0$, then $X_{it}^* = X_{it}$. If $X_{it} = 0$, then X_{it}^* is drawn from a truncated normal distribution with the mean $Z_{it}\delta + \frac{\sigma_{12}}{\sigma_{11}}\epsilon_{it}^Y$ and the variance $\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}$, where the truncation is to $(-\infty, 0)$.

4. $\delta|\beta_i, \alpha_i, \gamma, \Sigma_\epsilon, T, X^*$

Following Lahari and Schmidt (1978) and Rossi et al. (2005), we re-express the triangular system specified in Equation (1), (4) and (5) as a SUR model for making conditional draw of δ. Let $Y_i^* = Y_i - H_i(\beta_i, \alpha_i)'$, where H_i is given in Equation (15). We then have the likelihood of the triangular system the same as that of the following SUR model:

$$\begin{pmatrix} Y_{it}^* \\ X_{it}^* \end{pmatrix} = \begin{pmatrix} W_{it} & \\ & Z_{it} \end{pmatrix} \begin{pmatrix} \gamma \\ \delta \end{pmatrix} + \begin{pmatrix} \epsilon_{it}^Y \\ \epsilon_{it}^X \end{pmatrix}$$

where Y_{it}^* is the jth component of Y_i^*, W_{it} is the jth row in the design matrix W_i. Let $M_{it} = \begin{pmatrix} W_{it} & \\ & Z_{it} \end{pmatrix}$. Then (γ, δ) has a likelihood as that from a multivariate normal with the mean

$$(\bar{\gamma}, \bar{\delta}) = \left(\sum_{i,j} M_{it}' \Sigma_\epsilon^{-1} M_{it} \right)^{-1} \left(\sum_{i,j} M_{it}' \Sigma_\epsilon^{-1} \begin{pmatrix} Y_{it}^* \\ X_{it}^* \end{pmatrix} \right)$$

and the variance-covariance matrix

$$\begin{pmatrix} \Omega_{\gamma\gamma} & \Omega_{\gamma\delta} \\ \Omega_{\delta\gamma} & \Omega_{\delta\delta} \end{pmatrix} = \left(\sum_{i,j} M'_{it} \Sigma_{\epsilon}^{-1} M_{it} \right)^{-1}$$

Then given γ, the likelihood for δ is a multivariate normal with mean $\mu_{\delta|\gamma} = \bar{\delta} + \Omega_{\delta\gamma}\Omega_{\gamma\gamma}^{-1}(\gamma - \bar{\gamma})$ and variance-covariance matrix $\Omega_{\delta|\gamma} = \Omega_{\delta\delta} - \Omega_{\delta\gamma}\Omega_{\gamma\gamma}^{-1}\Omega_{\gamma\delta}$. Thus given the a normal prior $N(\mu_\delta, A_\delta^{-1})$ for δ, we can draw δ from a normal distribution with mean $(A_\delta + \Omega_{\delta|\gamma}^{-1})^{-1}(A_\delta\mu_\delta + \Omega_{\delta|\gamma}^{-1}\mu_{\delta|\gamma})$ and variance $(A_\delta + \Omega_{\delta|\gamma}^{-1})^{-1}$.

5. $\Sigma_\epsilon|\beta_i, \alpha_i, \gamma, \delta, T, X^*$

The conditional draw of Σ_ϵ follows $IW(\nu_\epsilon + N \times J, S_\epsilon + S)$, where $S = \sum_{i,j} \begin{pmatrix} \epsilon_{ij}^Y \\ \epsilon_{it}^X \end{pmatrix} (\epsilon_{it}^Y, \epsilon_{it}^X)$.

6. $\tau_i^*, \tau_i|\beta_i, \alpha_i, \gamma, \delta, \Sigma_\epsilon, X^*$

Random-Walk Metropolis-Hasting algorithm is used to update the individual-specific latent variable τ_i^* and τ_i. Draw a proposal $\tau_i^{*,prop}$ from $MVN(\tau_i^{*,old}, \kappa\Sigma_{e,22})$, where $\tau_i^{*,old}$ and $\Sigma_{e,22}$ are parameter draws at the previous iteration, and the scaling parameter κ is adjusted in the RWMH algorithm to achieve approximately 30% acceptance rate. Calculate $p(\tau_i^{*,prop}) = f_{\tau_i^*=\tau_i^{*,prop}}(Y_{it}|X_{it})f_\theta(\tau_i^{*,prop})$ and $p(\tau_i^{*,old}) = f_{\tau_i^*=\tau_i^{*,old}}(Y_{it}|X_{it})f_\theta(\tau_i^{*,old})$, where $f_{\tau_i^*}(Y_{it}|X_{it})$ is the likelihood function determined from Equations (14) and (15), and $f_\theta(\tau_i^{*,old})$ is the density function of prior distribution for τ_i. The prior distribution for τ_i^* is $N(\mu_{2|1,i}, A_{2|1}^{-1})$ and

$$\begin{aligned} \mu_{2|1,i} &= \mu_{2i} + \Sigma_{e,21}\Sigma_{e,11}^{-1}((\beta_i, \alpha_i)^T - \mu_{1i}) \\ A_{2|1}^{-1} &= \Sigma_{e,22} - \Sigma_{e,21}\Sigma_{e,11}^{-1}\Sigma_{e,12}, \end{aligned}$$

and μ_{1i} and μ_{2i} are the prior means of (β_i, α_i) and τ_i^*, respectively, and $(\mu_{1i}, \mu_{2i}) = \Pi Z_i$. Note that the above conditional mean and variance-covariance matrix should also condition on e^Z if LIVs are used for Z using the results derived in Step 7. Then with probability $q_i = min(1, p(\tau_i^{*,prop})/p(\tau_i^{*,old}))$, $\tau_i^{*,new} = \tau_i^{*,prop}$, and with probability $1 - q_i$, $\tau_i^{*,new} = \tau_i^{*,old}$. Given the draws of τ_i^*, the new draw of τ_i is determined by Equation (6).

7. Updating $\Pi, \Sigma_e|\beta_i, \alpha_i, Z_i$.

We first describe the algorithm for updating in the standard heterogeneity model where no LIVs are used to control for the potential endogeneity of Z. Let $\Lambda = \Sigma_e^{-1}$. We follow the standard approach (Gelman et al. 2004, Rossi et al. 2005) to obtain the conditional draws as follows:

$$p(\Lambda|\Theta) = \mathrm{W}\Big(\nu_e + N, \big(S_e + \sum_{i=1}^{N} e_i e_i'\big)^{-1}\Big),$$

where $e_i = B_i - \Pi Z_i$, $B_i = (\alpha_i, \beta_i, \tau_i^*)$, and

$$p(\Theta|\Lambda) = \mathrm{N}\Big(\Sigma_\Pi\big[(\Lambda \otimes I_{n_z})H_{zB} + \Lambda_\Pi\mu_\Pi\big], \Sigma_\Pi\Big),$$

N is the number of subjects, $\Sigma_\Pi = \big[(\Lambda \otimes H_{zz}) + \Lambda_\Pi\big]^{-1}$, and

$$H_{zz} = \sum_{i=1}^{N} Z_i Z_i', \qquad H_{z\beta} = \begin{bmatrix} H_{zB_1} \\ H_{zB_2} \\ \vdots \\ H_{zB_{n_r}} \end{bmatrix},$$

and $H_{zB_j} = \sum_{i=1}^{N} Z_i B_{it}$, for $j = 1, ..., n_r$.

We next move to the algorithm when LIVs are used to control for the potential endogeneity. For the most general case, because the system of equations expands to include models for Z as specified in Equations (13), the outcome vector is (B_i', Z_i'), we redefine Σ_e as the variance-covariance matrix of the expanded system, and $\Sigma_e = \begin{pmatrix} \Sigma_{e,BB} & \Sigma_{e,BZ} \\ \Sigma_{e,ZB} & \Sigma_{e,ZZ} \end{pmatrix}$.

• Update Π

Given the value of error terms in Z equations, e_i^Z, $B_i = (\alpha_i, \beta_i, \tau_i^*) = \Pi Z_i + e_i^B | e_i^Z$. Therefore, $B_i - \Sigma_{e,BZ}\Sigma_{e,ZZ}^{-1} e_i^Z = \Pi Z_i + e_i^{B|Z}$, where $e_i^{B|Z}$ is a conditional residual with mean zero and variance $\Sigma_{e,B|Z} = \Sigma_{BB} - \Sigma_{BZ}\Sigma_{ZZ}^{-1}\Sigma_{ZB}$. Because $e_i^{B|Z}$ is independent of Z_i, this reduces to the multivariate normal regression model and thus the standard algorithm described above for updating Π can be used to make conditional draw of Π.

• Update π

Consider the joint outcome $(B_i', Z_i')'$, we have

$$\begin{pmatrix} B_i \\ Z_i \end{pmatrix} = \begin{pmatrix} ZM_i & \\ & LM_i \end{pmatrix} \begin{pmatrix} \Theta \\ \pi \end{pmatrix} + \begin{pmatrix} e_i^B \\ e_i^Z \end{pmatrix} \tag{17}$$

where $ZM_i = \begin{pmatrix} Z_i' & & \\ & \cdots & \\ & & Z_i' \end{pmatrix}$, $LM_i = \begin{pmatrix} L_{i1}' & & \\ & \cdots & \\ & & L_{in_Z}' \end{pmatrix}$, and $\pi = (\pi_1', \cdots, \pi_{n_Z}')'$. Because the likelihood of the triangular system is the same as that of the SUR model (Lahari and Schmidt 1978), (Θ, π) has a likelihood as that from a multivariate normal with the mean

$$(\bar{\Theta}, \bar{\pi}) = \left(\sum_i M_i' \Sigma_e^{-1} M_i \right)^{-1} \left(\sum_i M_i' \Sigma_e^{-1} \begin{pmatrix} B_i \\ Z_i \end{pmatrix} \right)$$

and the variance-covariance matrix

$$\begin{pmatrix} \Omega_{\Theta\Theta} & \Omega_{\Theta\pi} \\ \Omega_{\pi\Theta} & \Omega_{\pi\pi} \end{pmatrix} = \left(\sum_i M_i' \Sigma_e^{-1} M_i \right)^{-1}$$

with $M_i = \begin{pmatrix} ZM_i & \\ & LM_i \end{pmatrix}$. Then given Θ, the likelihood for π is a multivariate normal with mean $\mu_{\pi|\Theta} = \bar{\pi} + \Omega_{\pi\Theta}\Omega_{\Theta\Theta}^{-1}(\Theta - \bar{\Theta})$ and variance-covariance matrix $\Omega_{\pi|\Theta} = \Omega_{\pi\pi} - \Omega_{\pi\Theta}\Omega_{\Theta\Theta}^{-1}\Omega_{\Theta\pi}$. Thus given the a normal prior $N(\mu_\pi, A_\pi^{-1})$ for π, we can draw π from a normal distribution with mean $(A_\pi + \Omega_{\pi|\Theta}^{-1})^{-1}(A_\pi \mu_\pi + \Omega_{\pi|\Theta}^{-1}\mu_{\pi|\Theta})$ and variance $(A_\pi + \Omega_{\pi|\Theta}^{-1})^{-1}$. To avoid label switching issue in Bayesian inference of LIV models, one approach is to impose restrictions on the parameters values of π. Let $Z_{ij} = \pi_{j0} + \pi_{j1} L_{ij} + e_{ij}^Z$, where we can impose restriction that $\pi_{j1} > 0$ to avoid label switching problem. One can then draw π_{j1} from the univariate truncated normal distribution derived from the above conditional distribution with the restriction $\pi_{j1} > 0$ (Geweke 1991).

• Update $L = (L_1, \cdots, L_{n_Z})$

The latent IV L_{ij} is draw from a Bernoulli distribution with the probability $p(L_{ij} = 1) \propto f(B_i, Z_i | L_{ij} = 1) p_j^L$, where $f(B_i, Z_i | L_{ij} = 1)$ is the density function of the SUR system given in Equation (17). The update is performed for all $i = 1, \cdots, N$ and $j = 1, \cdots, n_Z$.

• Update $p_L = (p_1^L, \cdots, p_{n_Z}^L)$

For each $j = 1, \cdots, n_Z$, the conditional draw of p_j^L follows a $Beta(1 + \sum_{i=1}^{N} L_{ij}, 1 + N - \sum_{i=1}^{N} L_{ij})$.

• Update Σ_e

The conditional draw of Σ_e follows $IW(\nu_e + N, S_e + S)$, where $S = \sum \begin{pmatrix} e_i^B \\ e_i^Z \end{pmatrix} (e_i^{B\prime}, e_i^{Z\prime})$.

References in Online Appendix

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2003), *Bayesian Data Analysis,* 2nd edition, Chapman & Hall, New York.

Geweke, J. (1991), *Efficient Simulation From the Multivariate Normal and Student t-Distributions Subject to Linear Constraints,* in *Computer Sciences and Statistics Proceedings of the 23rd Symposium on the Interface*, 571–578.

Lahari, K and Schmidt, P. (1978), "On the Estimation of Triangular Structural models," *Econometrica*,**46**, 1217–1221.

Rossi, P., Allenby, G.M. and McCulloch, R.E. (2005), *Bayesian Statistics and Marketing*, Hoboken, NJ: Wiley.

Web Appendix B: A Simulation Study

In this section, we conduct a set of simulation experiments to evaluate the performance of different models in repeated samples. We follow the steps below to simulate data:

- For each simulation dataset, we set the number of units $N = 30$ and the number of observations per unit $T = 12$, similar to those in the Chinese Shoe Market Data.
- To simulate data for Y_{it}, X^*_{it} and X_{it}, we first set the mean parameters in the following equations:

$$Y_{it} = \begin{cases} X_{it}\beta_{1i} + \alpha_i + W_{it}^T\gamma + \epsilon_{it}^Y & t = 1, ..., T_i - 1 \\ X_{it}\beta_{2i} + \alpha_i + W_{it}^T\gamma + \epsilon_{it}^Y & t = T_i, ..., T \end{cases}$$

and

$$X^*_{it} = \delta Z_{it} + \epsilon_{it}^X, \quad \text{and} \quad X_{it} = \begin{cases} X^*_{it} & \text{if } X^*_{it} \geq 0 \\ 0 & \text{if } X^*_{it} < 0. \end{cases}$$

In simulations, we set the short-term effect $\beta_1 = -0.4$ and the long-term effect $\beta_2 = 1.0$. The parameter vector γ includes the time fixed-effects that are simulated from $N(0, 0.1^2)$. The firm effects α_i is simulated from $N(0, 0.1^2)$. In simulations, we include in Z an intercept and three instrumental variables. The first IV is 0 before $J = 4$ and 1 afterward, mimicking the occurrence of a natural experiment. The second IV is simulated from N(0.5,1), and the third IV is the interaction of the first two IVs. We set the parameter values for coefficients on the intercept and the IVs to be $(-1, 2, 0, 0.5)$. We then randomly generated the response time τ_i^* from $N(2, 1.5^2)$ and form the grouped version of the response time τ_i according to Equation (6). The changepoint time $T_i = T_{ie} + \tau_i$, where T_{ie} is the first time that the treatment variable becomes positive.

- We generate the error terms $(\epsilon_{it}^X, \epsilon_{it}^Y)$ from a bivariate normal distribution with mean 0 and variance-covariance matrix

$$\Sigma_\epsilon = \begin{pmatrix} \sigma_{11} & \rho\sqrt{\sigma_{11}\sigma_{22}} \\ \rho\sqrt{\sigma_{11}\sigma_{22}} & \sigma_{22} \end{pmatrix},$$

where ρ is the correlation coefficient of two error terms. In simulations, we set $\sigma_{11} = 0.15^2$ and $\sigma_{22} = 0.5^2$, and we vary ρ in $(0, -0.2, 0.2, -0.5, 0.5, 0.8, -0.8)$ to cover various strengths of endogeneity in both directions.

- We then generate Y_{it}, X^*_{it} and X_{it} given the above model parameters. We repeat the above steps for $M = 50$ times for each parameter setting. This will generate 7*M panel datasets because there are seven values of ρ specified in step 3.
- We fit each simulated dataset with four models: the RC-SEM, SEM, RC model and the OLS model. Note that RC-SEM model is the full model and the other models can be considered as reduced models of RC-SEM. The RC model assumes $\rho = 0$ (i.e. no endogeneity issue) but allows heterogeneity of response times across units. In both SEM and the OLS the response times are assumed to be the same for all units. Furthermore, this common response time is not to be estimated from the data in SEM and OLS, but rather needs to be pre-specified. In the simulation study, we assume this common response time is 2, which is the mean of the response time used in the simulation. This corresponds to a scenario that the best guess of the response time, under the assumption of common response time, is used.
- We compare the estimates for both the population temporary short-term effect β_1 and the stable long-term entry effect β_2 from these four models. The result is summarized in Table 6. The columns "Bias", "SD" and "RMSE" are the bias, standard deviation, and square root of Mean squared error, respectively, calculated from the resulting sample of Bayesian estimates. We repeat the process for each value of ρ. The column "Coverage Rate" is the proportion of 95% credible intervals that contain the true values in the simulation experiments.

The simulation study shows that the estimation algorithm under the RC-SEM model recovers the true values of the temporary short-term β_1 and stable long-term entry effect β_2 reasonably well. Its RMSEs are smallest among four models across different strength of endogeneity. In addition, the coverage rates of the credible intervals are closest to the nominal 95% rate, among all methods. The simulation study shows that both the SEM model and OLS that ignores the heterogeneity in latent response times attenuate entry effects. The attenuation bias could be as large as 50% reduction in the true effect size. This shows that in the presence of heterogeneous response time, ignoring the heterogeneity and specifying a common response time can lead to severely biased effect estimates. Moreover, the Bayesian estimator from the RC-SEM model has less variability (i.e. smaller standard error) than that from the SEM, because the RC-SEM model provides better model-fitting by taking into account the latent response times. The RC estimates are biased because of the endogeneity issue. The OLS estimate has serious bias, particulary when the endogeneity is strong, and the 95% credible intervals hardly contain the true effect value.

Table 6 A simulation study on the comparison of performance of four models on estimating the entry effects. “Bias” and “SD” in the table are the bias and standard deviation of the estimates (posterior means), respectively, over all the replicates. “RMSE” denotes the root mean squared error. “CR” denotes the coverage rate.

ρ	RC-SEM				SEM				RC				OLS			
	Bias	SD	RMSE	CR	Bias	SD	RMSE	CR	Bias	SD	RMSE	CR	Bias	SD	RMSE	CR
Short-term effect (True value= -0.4)																
0	-0.01	0.12	0.12	92%	0.16	0.14	0.21	66%	-0.01	0.10	0.10	92%	0.16	0.13	0.20	72%
0.2	0.04	0.10	0.11	88%	0.22	0.13	0.25	52%	0.11	0.09	0.14	76%	0.27	0.12	0.30	32%
0.5	0.03	0.10	0.11	92%	0.25	0.15	0.29	46%	0.20	0.10	0.23	34%	0.38	0.16	0.41	18%
0.8	0.01	0.09	0.09	90%	0.25	0.14	0.27	42%	0.30	0.13	0.33	18%	0.45	0.16	0.48	8%
-0.2	0.01	0.10	0.10	98%	0.18	0.14	0.22	62%	-0.05	0.09	0.11	88%	0.12	0.13	0.18	78%
-0.5	0.02	0.10	0.10	92%	0.15	0.14	0.20	70%	-0.15	0.10	0.17	60%	0.03	0.14	0.15	84%
-0.8	0.00	0.09	0.09	94%	0.10	0.12	0.15	74%	-0.26	0.09	0.27	16%	-0.10	0.11	0.15	84%
Long-term effect (True Value=1.0)																
0	-0.01	0.08	0.08	94%	-0.16	0.10	0.18	54%	-0.01	0.06	0.07	94%	-0.17	0.09	0.18	46%
0.2	0.00	0.09	0.09	92%	-0.13	0.10	0.16	66%	0.07	0.09	0.12	74%	-0.07	0.10	0.12	80%
0.5	0.00	0.08	0.08	98%	-0.11	0.09	0.14	80%	0.18	0.08	0.20	24%	0.03	0.10	0.10	90%
0.8	-0.01	0.09	0.09	86%	-0.09	0.08	0.12	80%	0.30	0.10	0.31	4%	0.12	0.10	0.17	60%
-0.2	-0.01	0.09	0.09	92%	-0.18	0.11	0.21	50%	-0.09	0.07	0.12	80%	-0.25	0.10	0.27	16%
-0.5	0.01	0.09	0.09	88%	-0.19	0.10	0.21	46%	-0.17	0.08	0.19	36%	-0.33	0.10	0.34	6%
-0.8	-0.01	0.06	0.06	98%	-0.24	0.08	0.25	20%	-0.31	0.08	0.31	2%	-0.45	0.07	0.45	0%