

Gray Markets, A Product of Demand Uncertainty and Excess Inventory

Sriram Dasu[†], Reza Ahmadi[‡] and Scott M. Carr[‡]

[†] Marshall School of Business at USC, Los Angeles, CA 90089

dasu@marshall.usc.edu

[‡] UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA 90095-1481

reza.ahmadi@anderson.ucla.edu

Scott.Carr@anderson.ucla.edu

Gray Markets, A Product of Demand Uncertainty and Excess Inventory

Abstract

Diverting large quantities of goods from authorized distribution channels to unauthorized or “gray market” channels, albeit legal, significantly affects both firms and consumers due to effects on price, revenue, service and warranty availability, and product availability. In this paper we consider mechanisms by which the uncertainty surrounding inventory ordering decisions drives gray markets. We start with a minimal stochastic supply chain model composed of a producer and a retailer; then we restructure the model to add a distributor whereby the distributor and authorized retailer have the option of diverting inventory to a gray market. Our analysis sheds light on three issues: impacts of diversion on the various supply chain participants, strategies producers could use to combat or exploit gray markets, and important considerations for authorized retailers trying to set optimal order quantities in the presence of a gray market. Our analysis yields new insights into the behavior and impact of gray markets, which can inform management strategies and policies for confronting them.

Keywords: distribution channels, decisions under uncertainty, retailing and wholesale, gray markets

Originally submitted on January 2010

Revised on January 2011 and July 2011

Accepted on October 2011

1. Introduction

Information teased from a simple supply chain model can reveal factors leading to the legal diversion of products from authorized distribution and sales channels to gray markets, also called “unauthorized” or “parallel” channels, markets, or imports. Diversion occurs when a firm, such as a distributor or retailer, buys branded goods directly from the brand owner with the implied intention of selling the goods in one market (or country), but then actually selling them in a different one.

Large quantities of goods are diverted, rendering significant gray market impact. Estimates for unauthorized imports in the European Union include 10-20% of musical recordings, 5-10% of clothing, 5% of consumer electronics, up to 13% of cosmetics and perfumes, and up to 15% of soft drinks (NERA 1999 quoted in Maskus 2000 and AGMA 2004). The negative consequences of these unauthorized sales can be substantial for brand owners (Antia, Dutta, and Bergen 2004).

Consumers, on the other hand, can reap large positive as well as negative consequences. Though offering the promise of lower cost and improved product availability, unauthorized channels also carry the potential for service, compatibility, and warranty problems (Duhan and Sheffet 1988).

Distributors and other supply chain participants in gray markets often enjoy reduced inventory risk and significant revenue.

Pricing advantages are most commonly invoked to explain the existence of gray markets. As important and complex as pricing mechanisms are, however, other important factors also drive product diversion. In this paper, we assume set prices and focus on demand uncertainty. The realities of limited and random demand introduce the possibility that a supply-chain participant might possess more inventory than can be sold through authorized channels. In such situations, diverting excess inventory to an unauthorized channel, even one with lower margins than those in the authorized channel, might be common practice. Short life-cycle or seasonal products often

remain unsold in quantities as high as 25% of the initial stock, making significant quantities available for gray markets.

This paper describes a methodical exploration of the effects of demand uncertainty on gray market proliferation. Section 2 summarizes our investigation of the associated literature. With Section 3, we develop a parsimonious model consisting of a producer selling a short cycle fashion good directly to an authorized retailer; then, we add a second unauthorized channel providing a disposal option for retailers with surplus inventory. The unauthorized channel might consist of independent retailers in another market or a downstream wholesaler specializing in gray market items. We also consider a third channel, diverting inventory to salvage, for example donating excess clothing to charity as a tax write-off, selling excess glass as road-building material, or simply destroying excess product. Section 4 explores mechanisms for reducing diversions to gray markets. Section 5 extends our analysis to a more complex model in which a distributor buys from the producer and sells to both authorized and unauthorized retailers. The presence of a distributor diminishes the producer's ability to limit diversions, increases the supply of goods to gray markets, and lets authorized retailers continue to sell surplus goods to unauthorized retailers. Concluding remarks are provided in Section 6. Proofs are provided in the Appendix.

2. Literature review

The scenario is ubiquitous: a firm contributes to gray market activity by basing ordering decisions on uncertain demand. Most research into the scenario focuses on the role of price discrimination (Myers 1999, Michael 1998, Gerstner et al. 1994). Pricing issues have been investigated using game-theoretic models showing how unauthorized distribution channels provide opportunities for arbitrage and price discrimination. Readers are referred to Ahmadi and Yang (2000) and Ganslandt and Maskus (2004) and references therein. Chen (2009) finds that a parallel

gray market's impact on producer, authorized retailer, and customers depends on the relative price elasticity of demand.

A related area of research explores single period replenishment and transshipment decisions by multiple retailers or facilities owned by the same firm (Rudi, Kapur, and Pyke (2001), Dong and Durbin (2005), and Wee and Dada (2005)). Kouvelis and Gutierrez (1997) were among the first to study the impact of transshipments on production decisions. Our basic model can be viewed as a variant of theirs. However, we investigate the impact of such parameters as retail margins, demand uncertainty, and gray market size on producer profit and degree of diversion.

Finally, Lee and Whang (2002) study the impact of a secondary market on the profits of one producer and multiple identical buyers when the market arises from retailers selling surplus goods to one another at the end of the season. A closely related problem was also studied by Zhang and Chen (2003). Milner and Kouvelis (2007) extend this work by examining the relationship between long-term contracts and spot purchases in a multi-period context.

In all the models, initial purchase quantities are influenced by post-season opportunities; in ours, buyers sell their surplus inventory to unauthorized retailers. This gray market raises new questions. Can a gray market ever benefit the producer? What options are available to producers for mitigating the impact of gray markets? In the presence of gray markets, will growth in sales to the authorized channel result in proportional growth in the gray markets? Finally, how does the presence of an intermediary, such as a distributor, influence gray markets?

3. Impact of a gray market

In our model, a simplified supply chain consists of a producer and a retailer facing random demand for a pure fashion item. The item has such a short life-cycle that the retailer is presented with a single ordering opportunity. We also assume that the markets are competitive. Prices are exogenous and depend on the relative power of the various players (Betancourt and Gautschi 1993, Kadiyali,

Chintagunta, and Vilcassim 2000). Four links comprise the chain of events. (1) The authorized retailer (**a**) selects the quantity to order from the producer and receives the merchandise. (2) The selling season in the authorized sector occurs and resolves the demand uncertainty. (3) Any remaining inventory is then made available to the unauthorized channel (**u**). (4) Any inventory in excess of the unauthorized demand is then salvaged. Some notation is needed.

Channel	Inventory	Demand	Cost of Inventory	Sales price
Authorized	Q_a	D_a	C_1	P_a
Unauthorized	I_u^+	D_u	C_3	
Salvage	I_s^+		v	

Let quantities available to the secondary channel be unauthorized inventory $I_u^+ = \text{Max}(0, Q_a - D_a)$ and salvage inventory $I_s^+ = \text{Max}(0, Q_a - D_a - D_u)$. Authorized sales $D_a = Q_a - I_u^+$, unauthorized sales $D_u = I_u^+ - I_s^+$, and salvage sales are I_s^+ . The focus of this paper is to understand the issues that cause product to be diverted into the unauthorized channel, with the quantity diverted considered to be either unauthorized inventory, I_u^+ , unauthorized sales, $I_u^+ - I_s^+$, or even I_u itself. Under our assumptions, each of these three quantities increases if and only if the others do as well. We will thus use I_u^+ solely as the quantity diverted, and it will be the expected value $E[I_u^+]$ of unauthorized sales that is of greatest concern. The authorized retailer (**a**) will purchase the fashion item in a quantity Q_a intended to maximize profit. Thus,

$$Q_a^* = \text{Argmax}\{E[P_a(Q_a - I_u^+) + C_3(I_u^+ - I_s^+) + vI_s^+] - C_1Q_a\}.$$

A straightforward analysis shows that expected profit is a concave function of the order quantity. This requires additional notation.

Probability density function	of the demand (D_a) seen by...
$f_a(\cdot)$	Authorized retailer
$f_u(\cdot)$	Unauthorized retailer(s)
$f_{u a}(\cdot)$	Unauthorized retailer(s) conditioned on the demand observed by the authorized channel
$f_{a+u}(\cdot)$	Combined demand in both markets

The corresponding optimality condition is given by:

$$P_a(1 - F_a(Q_a^*)) + C_3 \int_{x=0}^{Q_a^*} [1 - F_{u|a}(Q_a^* - x)]f_a(x)dx + vF_a(Q_a^*) - C_1 = 0. \quad (1)$$

Note that the main difference between the gray market models and the transshipment models lies in the distribution of D_u , and $I_u^+ - I_s^+$ (Kouvelis and Gutierrez 1997, Lee and Whang 2002).

3.1. Unauthorized channel impact on authorized retailer purchase

Lacking the opportunity to sell to the gray market, the authorized retailer's optimal purchase quantity is the solution to the classical newsvendor problem given by $Q_N \equiv F_a^{-1}\left(\frac{P_a - C_1}{P_a - v}\right)$. If $C_3 > v$, then $Q_a^* \geq Q_N$ and the quantity purchased by the authorized retailer will increase. One immediate consequence of gray markets is that consumers in the authorized channel will enjoy greater service levels, but may see a decline in brand value due to the presence of the gray market.

3.2. Unauthorized channel impact on the producer

The prevailing sentiment in many industries is that unauthorized distribution channels harm producers (Antia, Dutta, and Bergen 2004). Diverted goods introduce significant warranty issues (Duhan and Sheffet 1988). Loss of brand value also may arise from the product being sold to segments that are not targeted by the producer. These costs are likely to increase with the size of the gray market. For the sake of generality and tractability, we model the negative consequences by introducing a linear penalty that is incurred by the producer for each unit sold in an unauthorized channel.

To investigate conditions producing the greatest penalties for diversion, we require added notation.

π	Penalty > 0 incurred by the producer	σ_a	Uncertainty in authorized market
λ_a	Size of authorized market	σ_u	Uncertainty in unauthorized market
λ_u	Size of unauthorized market	ϕ	Probability density function
		Φ	Cumulative distribution function
M	Unit margin earned by the producer for each unit sold to the authorized retailer	α_1	Service level in authorized market
		α_2	Service level in combined markets
		ρ	Correlation coefficient

When margins in the authorized markets are high, levels of I_u^+ will be high and, thus, the quantity diverted to the gray markets will be high. At the same time, the gray market will induce only a small amount of additional purchases by the authorized retailer. Therefore, gray markets are likely to be most harmful to producers in markets where authorized service levels are high, so that the retailer maintains considerably more stock on hand than consumers using the authorized channel demand. To gain further insight, we assume that demand in these two markets displays a bivariate normal distribution with a correlation coefficient ρ . That is, the authorized demand is $N(\lambda_a, \sigma_a^2)$, unauthorized demand is $N(\lambda_u, \sigma_u^2)$, and the total for both is $N(\lambda_a + \lambda_u, \sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u)$. From known results for truncated distributions (Patel and Read 1982), expected quantities diverted and salvaged are clearly:

$$E[I_u^+] = (Q_a - \lambda_a) \Phi\left(\frac{Q_a - \lambda_a}{\sigma_a}\right) + \sigma_a \phi\left(\frac{Q_a - \lambda_a}{\sigma_a}\right) \text{ and}$$

$$E[I_s^+] = (Q_a - \lambda_a - \lambda_u) \Phi\left(\frac{Q_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right) + \sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u} \phi\left(\frac{Q_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right). \quad (2)$$

For ease of exposition, let α_1 denote $\Phi\left(\frac{Q_a - \lambda_a}{\sigma_a}\right)$, the service level in the authorized market, and α_2 denote $\Phi\left(\frac{Q_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right)$, the service level in the combined market. The producer's penalty is:

$$\pi \cdot E[I_u^+ - I_s^+] = \pi \left[(Q_a - \lambda_a)(\alpha_1 - \alpha_2) + \alpha_2 \lambda_u + \sigma_a \phi\left(\frac{Q_a - \lambda_a}{\sigma_a}\right) - \sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u} \phi\left(\frac{Q_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right) \right]. \quad (3)$$

Recall that the expected unauthorized sales are $E[I_u^+ - I_s^+]$. Equation (3) suggests that the penalty may increase if the service level, α_1 , increases; the size of unauthorized market, λ_u , increases; the uncertainty in the authorized market, σ_a , increases; or the uncertainty in the unauthorized market, σ_u , decreases. High service levels in the authorized market increase the

amount available at the end of the authorized season. The excess at the end of the season as a

percentage of the optimal purchase quantity, Q_a , is given by:
$$\frac{E[I_u^+]}{Q_a} = \frac{\frac{\sigma_a}{\lambda_a}(\alpha_1 \Phi^{-1}(\alpha_1) + \phi(\Phi^{-1}(\alpha_1)))}{(1 + \Phi^{-1}(\alpha_1) \frac{\sigma_a}{\lambda_a})}$$
.

This ratio is a function of the coefficient of variation, $\frac{\sigma_a}{\lambda_a}$, and of the optimal service level, α_1 .

We find that when the coefficient of variation is as low as 0.25 and service levels are greater than 0.7, then an excess of 15% of the initial purchases will be available through gray markets. Thus, the quantities that can be diverted are quite significant.

3.3. Authorized and unauthorized market size and profit margin impact on quantity diverted

Proposition 1. *The expected quantity diverted, $E[I_u^+]$, to the unauthorized market (i) is constant for expected authorized demand, λ_a , (ii) increases with expected unauthorized demand, λ_u , and (iii) increases with the margins of all three channels.*

Of these, (i) is the most interesting. It is intuitive and also true that, for any given order size, an increase in authorized demand will reduce the quantity diverted. At first glance, this seems to contradict (i), which says that quantity diverted will remain constant. The proposition holds, however, because an increase in λ_a will induce the authorized retailer to change her order quantity by the exact amount that λ_a changes. The importance of (i) is that producers, who naturally use advertising and other methods to generate additional demand, need not be concerned that they are also inducing additional diversion and its attendant problems. A result analogous to (iii) is observed in secondary markets. Lee and Whang (2002) find, when the margins are high, initial purchase quantities decline and the quantity of trade in the secondary markets increase.

The following proposition partially characterizes the effects of demand uncertainty on diversion quantities. A more complete description is greatly complicated by the intractability of the standard normal cdf, Φ , that enters several of the underlying derivatives.

Proposition 2. *The expected quantity diverted, $E[I_u^+]$, to the unauthorized market (i) increases (decreases) with unauthorized demand uncertainty if and only if $Q_a > (<) \lambda_a + \lambda_u$, and (ii) increases with authorized demand uncertainty if $Q_a > \lambda_a + \lambda_u$.*

Demand uncertainty impact depends on the service levels of the combined market. If the service level exceeds 50%, then an increase in demand uncertainty increases the expected quantity available for the gray markets. When the service level is above (below) 50% the optimal purchase quantity exceeds (is less than) the mean. If the standard deviation increases, then the absolute difference between optimal purchase quantity and average demand also increases.

3.4. Impact of incremental sales to the authorized channel

We now consider the effects of additional purchases and the dependence of producer profit on retail ordering. The producer's expected benefit from sales to the retailer, B_p , is:

$$B_p = MQ_a - \pi \cdot E[I_u^+ - I_s^+]. \quad (4)$$

By studying $\frac{\partial B_p}{\partial Q_a}$, we gain insight into the marginal change in producer benefit due to changes in purchase quantity.

Proposition 3. *The producer's profits increase with the quantity of items purchased by the authorized retailer; i.e. $\left(\frac{\partial B_p}{\partial Q_a}\right) > 0$, if and only if:*

$$\frac{M}{\pi} > F_a(Q_a) - F_{a+u}(Q_a). \quad (5)$$

Because the right side of the Equation (5) never exceeds one, even when the penalty, π , exceeds the margin, M , the producers might benefit from an increase in order quantity. The producer has little to fear from increased order quantities if his margin is nearly as large as the penalty. In fact, when the initial service level in the authorized channel is low, there is a high probability that additional purchases by the authorized retailer will be sold in the authorized channel. Additional sales generate $M(Q_a - Q_N)$ in additional profits, and sales in the unauthorized channel produce a penalty of

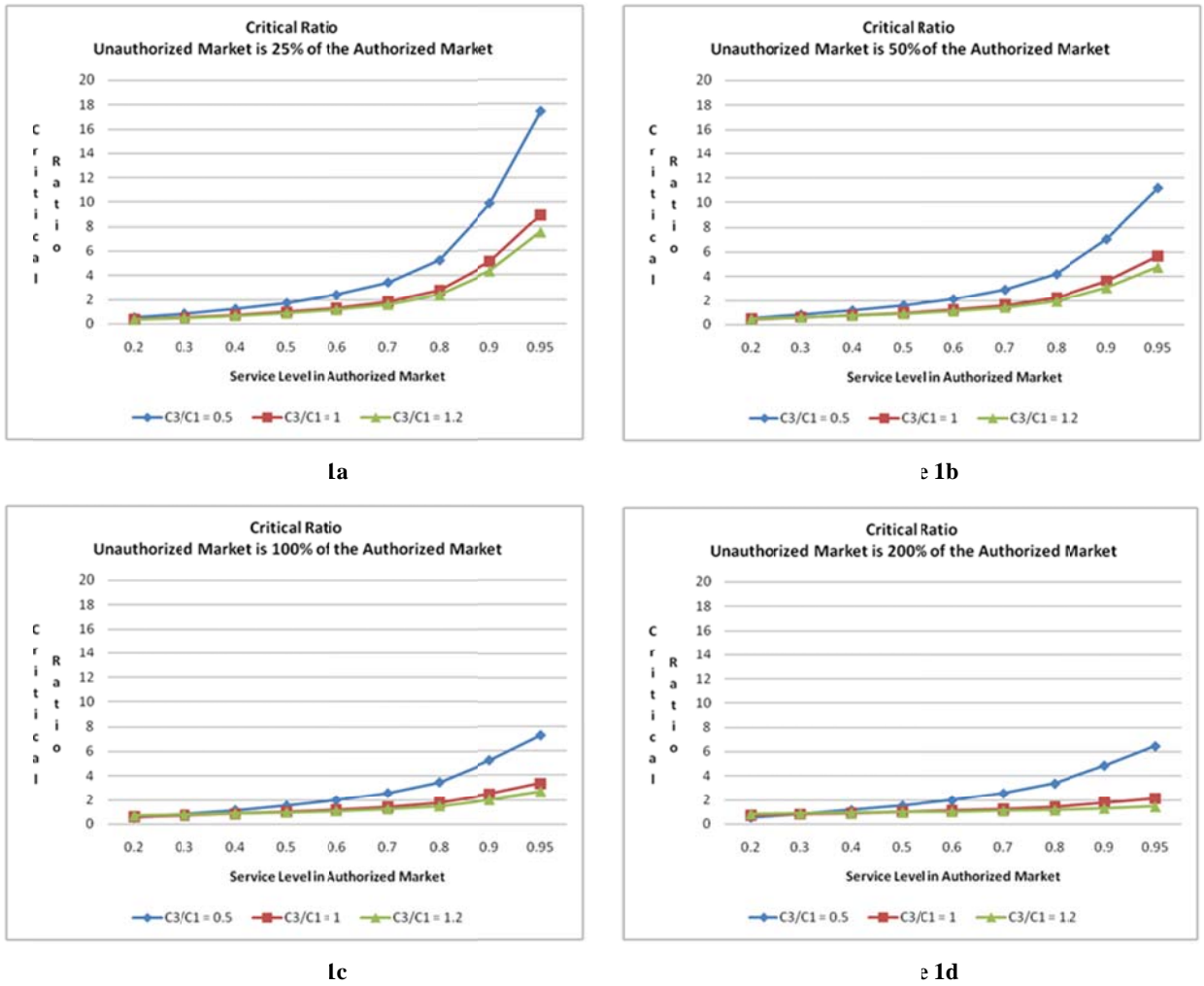
$\pi E[I_u^+ - I_s^+]$. The producer benefits from a gray market as long as $\frac{M}{\pi} > \frac{E[I_u^+ - I_s^+]}{Q_a - Q_N}$, which we refer to as the *critical ratio*. The primary harm to the producer, therefore, probably comes when goods move from the salvage market to the gray market.

3.5. Impact of various parameters on the critical ratio

The gray market cost to the producer increases when the critical ratio increases. Parameters influencing the ratio may include relative sizes of the authorized and unauthorized markets, service level in the authorized market, and ratio of inventory cost for the unauthorized retailer to that for the authorized retailer, (C_3/C_1) . When C_3/C_1 is less than one, then the authorized retailer sells to the gray market at a loss. In Figure 1, we graph the relationship between service level and inventory cost ratio using four relative sizes of unauthorized market.

Counter to what one might expect, Figure 1 suggests that the critical ratio can be smaller for larger gray markets and for higher gray market inventory costs, C_3 . An increase in the size of the gray market or an increase in the inventory cost induces the authorized retailer to purchase additional units from the producer. From Proposition 3, we know that additional purchases do not harm the producer much, because there is a chance that these units will be sold in the authorized market. Most of the harm to the producer is due to the initial surplus that exists even in the absence of gray markets. Consistent with this, we find that the critical ratio increases with service level.

Figure 1. Relationship between critical ratio and service level for four unauthorized market sizes



Two examples illustrate the impact of gray market size on producer earnings and provide details showing how earnings change relative to the graphs. In one, retail margins and service levels are high, and in the other, both are low. The authorized sales price $P_a = 100$ for both examples.

Example: High retail margins and low penalty. In this case, $C_3/C_1 = 0.5$ and the penalty from diversions is small relative to producer margins, making the producer worse off if the gray market is small. Authorized demand is $N(200,60)$, and gray market demand is also normal with a coefficient of variation of 0.15. These markets have a correlation of 0.4. The producer's cost is \$4/unit, $C_1 = \$20$ /unit, $C_3 = \$10$ /unit, $v = \$0$ /unit, $M = \$16$ /unit, and $\pi = \$4.25$ /unit. The quantity purchased by the authorized retailer, the expected sales in the gray market, and the producer's

profits for different sizes of the gray market are shown in Table 2. Observe that if the average demand in the unauthorized market is less than 200 units, then the gray market decreases the producer's net earnings. If the demand is higher than 200 units, the producer is better off with a gray market.

Table 2 Impact of gray markets on producer's profits (high service levels)

λ_u	0	50	100	200	400
Q_N	251				
Q_A		257	264	272	273
Expected gray market sales		33.53	55.64	73.46	76.41
Earnings from sales	\$4,008	\$4,109	\$4,220	\$4,345	\$4,371
Penalty	\$0.00	\$143	\$236	\$312	\$325
Net earnings	\$4,008	\$3,966	\$3,984	\$4,033	\$4,046

Example: Low retail margins and high penalty. Here, demand distributions are the same as those in the previous example, but $C_3/C_1 = 1.2$. Producer cost is \$4/unit, $C_1 = \$80$ /unit, $C_3 = \$96$ /unit, $v = \$0$ /unit, $M = \$76$ /unit, and $\pi = \$152$ /unit. The quantity purchased by the authorized retailer, the expected sales in the gray market, and producer profits for different gray market sizes are shown in Table 3. Although the penalty is twice the margin, observe that earnings are higher when λ_u is less than 100. In this example, the gray market causes the authorized retailer to buy additional goods. Initially the goods flow to the authorized channel, but, eventually, the bulk of additional goods is diverted to the gray market.

Table 3 Impact of gray markets on producer's profits (low service levels)

λ_u	0	50	100	200	400
Q_N	149				
Q_A		194	237	326	504
Expected gray market sales		14.41	40.46	119.64	294.56
Earnings from sales	\$11,324	\$14,719	\$18,007	\$24,793	\$38,271
Penalty	\$0	\$2,190	\$6,150	\$18,185	\$44,773
Net earnings	\$11,324	\$12,529	\$11,857	\$6,607	-\$6,501

These examples illustrate the complex relationship between gray markets and producer profit. If the penalty is low, large gray markets are preferable to small gray markets. Interestingly,

however, even when penalties are large, gray markets can be beneficial, provided initial service levels are low and gray markets are small. This suggests that, if the producer can influence price in the presence of gray markets, he should raise C_1 , his price to the authorized retailer.

4. Mechanisms to reduce gray markets impact

Clearly, gray markets diminish producer profit unless margins are significantly larger than the penalty. An obvious mitigation tactic is to change the inventory cost to authorized retailers, but this paper focuses on situations in which the producer has limited ability to influence prices and must employ other mechanisms. We consider buyback contracts and multiple replenishments.

4.1. Buyback contracts

In a buyback contract (Cachon 2003), the producer sells to the retailer at a price, w , and repurchases any surplus at a price, b . In the classical newsvendor setting, w and b are chosen to maximize channel profits. Greater values of b reduce retailer profit and risk. In the presence of gray markets, the producer will want b to be greater than C_3 in order to ensure that the authorized retailer returns excess inventory to the producer. By participating in a buyback arrangement with a greater value of C_3 , authorized retailer profit will decline and producer profit will increase. In the absence of buyback contracts, the opposite occurs: higher C_3 values return a larger profit to authorized retailers.

A measure of product appeal. The price that the unauthorized retailer is willing to pay indicates the appeal of the product. Higher levels of C_3 should increase the power of the producer and thereby cause C_3 to be the floor for b . However, unless the authorized retailer enjoys consistently high profit selling only in the authorized channel, she may not agree to a buyback contract.

Example: Limitations of buybacks. We examined a conservative situation in which $P_a = \$100$, $C_0 = \$10$, $C_1 = \$30$, $C_3 = \$36$, $v = \$0$, authorized channel demand is $N(100,30)$, gray market demand is $N(200,60)$, and $\rho = 0.4$. By selling to the gray market, the retailer earns \$7,457. If the buyback price b is \$36, channel profit is maximized at $w = \$42.4$, and the retailer earns \$5,423. Even if the producer lowers w to \$36.1, the retailer earns only \$6,397, suggesting that there is no acceptable buyback contract.

4.2. Multiple replenishments

Because sales in the early part of the season inform total sales for the season (Fisher and Raman 1996), ongoing replenishments should be based on forecasts with lower levels of uncertainty and should bring the total quantity purchased closer to the true demand, thus reducing the quantity supplied to the unauthorized channel. To explore the validity of this intuition, we will study a system in which shortages during the first half of the season result in lost sales, and the producer replenishes stocks midway, as depicted in Figure 2.

Figure 2. Sequence of events whereby the authorized retailer purchases twice in a season



The value of two replenishments depends on retail margins, size of the gray market, and price to the gray market, C_3 . Proposition 4 gives us the means to compare performance of the dual replenishment system to that of a single replenishment system for different prices of C_3 and different service levels. Additional notation:

$Q_{1,a}$	Quantity purchased at the beginning of the season
d_1	Sales in the first half of the season
$Q_{2,a}(d_1)$	Quantity purchased in the middle of the season (order-up-to level)

$f_{1,a}(\cdot)$ Distribution of demand in the first half of the season
 $f_{2,a}(\cdot | d_1)$ Distribution of demand in the second half of the season

Proposition 4. *Optimal purchase quantities $Q_{1,a}$ and $Q_{2,a}(d_1)$ are determined as follows. Let*

$Q_{2,a}^(d_1)$ be the unique solution to:*

$$P_a \int_{x=Q_{2,a}^*(d_1)}^{\infty} f_{2,a}(x|d_1) dx + C_3 \int_{x=0}^{Q_{2,a}^*(d_1)} \int_{y=Q_{2,a}^*(d_1)-x}^{\infty} f_u(y|x+d_1) f_{2,a}(x|d_1) dx + v \int_{x=0}^{Q_{2,a}^*(d_1)} \int_{y=0}^{Q_{2,a}^*(d_1)-x} f_u(y|x+d_1) f_{2,a}(x|d_1) dx - C_1 = 0 \quad (6)$$

$$Q_{2,a}(d_1) = \text{Max}(0, Q_{2,a}^*(d_1) - Q_{1,a} - d_1). \quad (7)$$

Let $Q_{1,a}$ be the unique solution to:

$$P_a \int_{x=Q_{1,a}}^{\infty} f_{2,a}(x|d_1) dx + C_1 \int_{x=x^*}^{Q_{1,a}} f_{1,a}(x) dx + P_a \int_{x=0}^{x^*} \int_{y=Q_{1,a}-x}^{\infty} f_{2,a}(y|x) f_{1,a}(x) dy dx + C_3 \int_{x=0}^{x^*} \int_{y=0}^{Q_{1,a}-x} \int_{z=Q_{1,a}-x-y}^{\infty} f_u(z|y+x) f_{2,a}(y|x) f_{1,a}(x) dz dy dx + v \int_{x=0}^{x^*} \int_{y=0}^{Q_{1,a}-x} \int_{z=0}^{Q_{1,a}-x-y} f_u(z|y+x) f_{2,a}(y|x) f_{1,a}(x) dz dy dx - C_1 = 0 \quad (8)$$

In which

$$x^* = \text{sup}(x: Q_{1,a} - Q_{2,a}(x) > 0). \blacksquare \quad (9)$$

Impact on authorized retailer purchases: The authorized retailers expected purchase quantity increases when there are two purchases as shown in Table 4. We assumed that demand in the two halves is positively correlated, make the authorized sales price $P_a = 100$, and set the salvage value $v = 0$. Details of the demand distributions can be found in the Appendix. The optimization problem for the second purchase is identical in structure to that faced by a retailer making a single purchase. At the time of the second purchase, however, demand uncertainty is lower and the gray market is larger relative to the authorized demand. Consequently, the second order-up-to level, $Q_{2,a}^*(d_1)$, will cover a greater percentage of authorized demand compared to the percentage of demand covered by the retailer who makes a single purchase.

Table 4. Percent change in purchase amounts with two replenishments

Service Level ($P_a - C_1$)/ P_a	C_3/C_1		
	0.5	1	1.2
0.2	24.71%	31.00%	36.42%
0.3	21.20%	27.87%	33.44%
0.4	19.21%	26.13%	31.16%
0.5	17.59%	23.92%	28.93%
0.6	16.19%	22.12%	26.89%
0.7	14.89%	20.41%	24.45%
0.8	13.51%	18.53%	22.01%
0.9	11.71%	15.81%	18.58%
0.95	10.44%	13.87%	16.02%

As for the first purchase of a retailer making two purchases, any excess from the first purchase will be used in the second period or will be sold on the gray market. Hence, the effective salvage value will increase, the initial purchase will cover a greater percentage of first-half demand compared to the percentage of total demand covered by the retailer who makes a single purchase for the entire season, and total expected purchases by the authorized retailer will increase.

Impact on expected quantity diverted to the gray market. Table 5 shows that the quantity diverted may rise, so that, in certain situations, offering multiple replenishments harms the producer. However, diversions to the gray market decrease when $C_3 < C_1$.

Table 5. Percentage increase in quantity sold to the gray market with two replenishments

Service Level ($P_a - C_1$)/ P_a	C_3/C_1		
	0.5	1	1.2
0.2	-21.98%	20.78%	49.67%
0.3	-26.85%	9.29%	34.72%
0.4	-28.80%	3.45%	24.69%
0.5	-29.79%	-2.19%	16.47%
0.6	-30.21%	-6.61%	9.48%
0.7	-30.18%	-10.18%	2.87%
0.8	-29.66%	-13.40%	-3.19%
0.9	-28.24%	-16.37%	-9.44%
0.95	-26.41%	-17.56%	-12.70%

When the authorized retailer finds it unattractive to sell to the gray market, multiple replenishments diminish gray market sales. On the other hand, if the gray market is attractive and service levels are low, the same retailer will supply the gray market by exploiting her enhanced ability to match supply and demand. In summary, multiple replenishments help producers when margins are high in the authorized channel, but, when margins are low and gray market prices are high, multiple replenishments turn authorized retailers into distributors, thus harming producers.

5. Impacts of a gray market with intermediation

In this section, we augment our model with a distributor that intermediates between both types of retailers. The distributor supplies goods to the unauthorized channel at the beginning of the season, a situation observed in retail sectors of some emerging markets. Here, big authorized retailers serve high income segments and small unauthorized retailers serve low income segments. At the end of his selling season, the authorized retailer sells his surplus to unauthorized retailers, whose selling season extends longer, or salvages it. This sequence is depicted in Figure 3, with Period 1 being the authorized selling season and Period 2 being the extended unauthorized selling season.

Figure 3. Sequence of events in a channel with a distributor and two classes of retailers



5.1. Optimal order quantities for authorized and unauthorized retailers

In order to determine the optimal order quantity and expected profit for each of the retailers, we make a number of assumptions that simplify the analysis without altering the potential for qualitative insights. We assume that the smaller retailers are identical and that they do not compete with each other. Because the surplus available to the authorized retailer, **a**, at the end of Period 1 may not be adequate to meet the demand from all of the smaller retailers, we assume a proportional rationing scheme, which allows us to treat the smaller retailers collectively as one retailer, **u**. We also assume that by the end of Period 1, **u** can accurately predict the total demand. This simplification is partially justified by prior work suggesting that early demand information is a powerful predictor of total season sales (Fisher and Raman 1996).

For **u**, let θ , be the season overlap with that of **a**. In other words, demand in Period 1 is a fraction, θ , of the total demand noted by **u**. The ensuing analysis shows how **a** and **u** can incorporate each other's decisions while optimizing their purchases. We require added notation:

C_2	Price at which the distributor sells to u at the beginning of the season
Q_u	Quantity u buys from the distributor at the beginning of the season
Z	Surplus held by a at the end of his selling season
Y	Quantity u is willing to purchase from a
$\Pi_a(Q_a, Q_u)$	Expected profits of a as a function of Q_a and Q_u
$\Pi_u(Q_a, Q_u)$	Expected profits of u as a function of Q_a and Q_u

At the beginning of Period 1, **a** and **u** make purchasing decisions informed by the possibility that **a** would sell his surplus to **u**. If **a** purchases Q_a units, intending to sell surplus to **u**, the quantity

\mathbf{u} would buy from \mathbf{a} depends on Q_u and demand, D_u . We expect the demand observed by \mathbf{u} and \mathbf{a} to be related; therefore, the quantity that \mathbf{u} buys from \mathbf{a} also depends on D_a . Let $f_a(D_a|D_u)$, and $f_u(D_u|D_a)$ denote the conditional demand distributions for \mathbf{a} and \mathbf{u} , respectively.

Let $f_Y(y|D_a, Q_u)$ denote the conditional distribution of Y . We assume that \mathbf{u} will purchase from \mathbf{a} only to meet demand in the second period and that shortages in Period 1 are not backlogged. If $D_u \leq Q_u$, then $Y = 0$. If $Q_u \leq D_u \leq Q_u/\theta$, \mathbf{u} will be able to meet demand in Period 1 and will require $D_u - Q_u$ units to meet demand in Period 2, so that $Y = D_u - Q_u$. If $D_u \geq Q_u/\theta$, then \mathbf{u} runs out of stock in Period 1. She will purchase $(1 - \theta)D_u$ to meet her requirements in Period 2, thus:

$$Pr_Y(y = 0|D_a, Q_u) = \int_{s=0}^{Q_u} f_u(s|D_a) ds, f_Y(y|D_a, Q_u) = f_u(y + Q_u|D_a) \text{ for } 0 < y \leq \frac{(1-\theta)}{\theta} Q_u, \text{ and } f_Y(y|D_a, Q_u) = f_s\left(\frac{y}{1-\theta}|D_a\right) \text{ for } \frac{(1-\theta)}{\theta} Q_u \leq y.$$

Let $f_Z(z|D_u, Q_a)$ denote the conditional distribution \mathbf{a} 's surplus held by at the end of Period 1, so:

$$Pr_Z(z = 0|D_u, Q_a) = \int_{r=0}^{Q_a} f_a(r|D_u) dr \text{ and } f_Z(z|D_u, Q_a) = f_a(Q_a - z|D_u, Q_a) \text{ for } 0 < z \leq Q_a.$$

Now let $L_a = \operatorname{argmax}_r: F_a(r|D_u, Q_a) = 0$ and $H_a = \operatorname{argmin}_r: F_a(r|D_u, Q_a) = 1$, in which $F(\cdot)$ denotes the cumulative demand distributions. If we also let L_u and H_u represent corresponding variables for $F_u(s|D_a, Q_u)$, we can make the following benign assumption about the demand distributions.

Assumption For $L_a \leq r \leq H_a$ $f_a(r|D_u, Q_a) > 0$, and for $L_u \leq r \leq H_u$ $f_u(s|D_a, Q_u) > 0$.

Given purchase decisions Q_a and Q_u , we need to consider two cases in order to calculate \mathbf{a} 's expected profit. First, if $Q_a \leq D_a$, then \mathbf{a} will sell all Q_a at price P_a to his customers. Second, if $Q_a \geq D_a$, then \mathbf{a} will sell D_a units at a price P_a , sell $\max(Q_a - D_a, Y)$ at price C_3 to \mathbf{u} , and salvage the rest at a price v . Combining the cases we get the following expression for \mathbf{a} 's expected profit:

$$\begin{aligned} \Pi_a(Q_a, Q_u) = & \int_{r=0}^{Q_a} \left\{ P_a r + \int_{y=0}^{Q_a-r} (C_3 y + V(Q_a - r - y)) f_y(y|r, Q_u) dy + \int_{y=Q_a-r}^{\infty} C_3 (Q_a - r) f_y(y|r, Q_u) dy \right\} f_R(r) dr + \\ & \int_{r=0}^{Q_a} v Pr_Y(0|r, Q_u) (Q_a - r) f_R(r) dr + \int_{r=Q_a}^{\infty} P_a Q_a f_R(r) dr - C_1 Q_a. \end{aligned} \quad (12)$$

Proposition 5. Let $G_a(Q_u) = \text{Argmax}_Q \Pi_a(Q, Q_u)$. $\Pi_a(\cdot, Q_u)$ is a strictly convex function and $G_a(Q_u)$ is unique.

To determine \mathbf{u} 's expected profit, we need to consider three more cases. First, if $\theta D_u \geq Q_u$, \mathbf{u} will run out of stock before Period 1 ends. Thus \mathbf{u} will sell Q_u units during \mathbf{a} 's selling season and will try to purchase $(1 - \theta)D_a$ units from \mathbf{a} . If $Q_u \leq D_u \leq Q_u/\theta$, the expected profit in is given by:

$$\int_{s=\frac{Q_u}{\theta}}^{\infty} \left\{ P_u Q_u + \int_{z=0}^{(1-\theta)s} (P_u - C_3) z f(z|s, Q_a) dz + \int_{z=(1-\theta)s}^{\infty} (P_u - C_3) (1 - \theta) s f(z|s, Q_a) dz \right\} f_s(s) ds. \quad (13)$$

Second, \mathbf{u} has enough stock to meet demand during the selling season for \mathbf{a} , but wants to buy $D_u - Q_u$ additional units from \mathbf{a} . The actual purchase quantity will depend on \mathbf{a} 's surplus, Z . The expected profit in this case is given by:

$$\int_{s=\frac{Q_u}{\theta}}^{\frac{Q_u}{\theta}} \left\{ P_u \theta s + \int_{z=0}^{s-Q_u} (P_u(z + Q_u - \theta s) - C_3 z) f(z|s, Q_a) dz + \int_{z=s-Q_u}^{\infty} (P_u(1 - \theta)s - C_3(s - Q_u)) f(z|s, Q_a) dz \right\} f_s(s) ds + \int_{s=\frac{Q_u}{\theta}}^{\frac{Q_u}{\theta}} P_u(Q_u - \theta s) Pr_Z(0|s, Q_a) f_u(s) ds. \quad (14)$$

Third, if $D_u \geq Q_u$, then \mathbf{u} will not buy from \mathbf{a} and will have to salvage the surplus. The expected profits are given by:

$$\int_{s=0}^{Q_u} \{ P_u s + v(Q_u - s) \} f_s(s) ds. \quad (15)$$

Therefore,

$$\Pi_u(Q_a, Q_u) = (13) + (14) + (15) - C_2 Q_u. \quad (16)$$

Proposition 6. The objective $\Pi_u(Q_a, \cdot)$ is concave in Q_u if we fix Q_a . Let $G_u(Q_a) = \text{Argmax}_Q \Pi_u(Q_a, Q)$ then $G_u(Q_a)$ is the unique optimal solution.

Proposition 7. $G_u(\cdot)$ and $G_a(\cdot)$ are monotone non-increasing functions.

Proposition 8. If $G_u(\cdot)$ and $G_a(\cdot)$ are continuous functions, there exists an equilibrium solution.

5.2. Influence of costs and margins on the flow of goods

The complexity of the model makes it very difficult to elicit insights using analytics, so we will resort to numerical experiments. The demand for the two markets has a correlation of 0.5. Demand for the authorized channel is $N(200,60)$. The coefficient of variation of demand in the gray market is 0.15. In all our examples, the equilibrium is unique. We also assume the following:

- eighty percent of gray market sales occur after the authorized retailer's season ends;
- selling price in the gray market is 80% of the regular market price; and
- salvage value $v = 0$ for the authorized and the unauthorized retailer.

The percentage of gray market inventory due to diversions by the authorized retailer is shown in Table 6. In these examples, only 20% of gray market sales occur during the authorized selling season. Furthermore, gray market sales during the second period are known with certainty at the end of the first period; yet, the bulk of the gray market supply comes from the distributor.

Influence of margins on unauthorized retailer purchasing decisions. The unauthorized retailer must determine her purchase quantity while taking demand and supply uncertainties into account. When retail margins are high, the unauthorized retailer is willing to absorb the risk of demand uncertainty and purchase the bulk of her requirements from the distributor. The ratio $(P_a - C_1)/P_a$ is a measure of the retail margins. When retail margins are low, the unauthorized retailer is willing to wait until the demand uncertainty is resolved. In this case, she will buy more of her goods from the authorized retailer. When the authorized retailer purchases directly from the producer, the supply to the unauthorized sector is monotone, increasing in retail margins. The presence of a distributor reverses the pattern. Table 6 shows that the gray market purchases a large percentage from the authorized retailer. In this example, C_3/C_1 is 0.5 and C_2/C_1 is 1.2. Thus, the distributor is charging the gray market a premium, while the authorized retailer is selling to the gray market at a price lower than his purchase price. For these costs, the unauthorized retailer strongly

prefers to delay her purchases when margins are low. In all other cases, she makes the bulk of her purchases at the beginning of the season from the distributor.

Table 6. Gray market sales by the authorized retailer as a percentage of total gray market supply

#	λ_u/λ_a	C_3/C_1	C_2/C_1	$(P_a - C_1)/P_a$					
				0.4	0.5	0.6	0.7	0.8	0.9
1	0.5	0.5	1.2	52.31%	15.44%	9.79%	7.59%	6.19%	4.67%
2	0.5	1.2	0.5	0.49%	0.55%	0.59%	0.60%	0.62%	0.54%
3	0.5	1.2	1.2	7.54%	5.49%	4.61%	4.11%	3.58%	2.89%
4	0.5	0.5	1.0	5.67%	4.99%	4.56%	4.14%	3.65%	2.93%
5	0.5	1.0	1.0	3.70%	3.39%	3.20%	2.95%	2.68%	2.26%
6	0.5	1.2	1.0	3.14%	2.91%	2.75%	2.57%	2.34%	1.98%

Effect on unauthorized sales due to being supplied by the authorized channel. In Table 7, we compare the quantity sold in the gray market when the authorized channel does not divert to the quantities when it does diverts. The results are counter-intuitive. If the authorized retailer lowers the price to the unauthorized channel, gray market sales decrease. When the price drops below that at which the distributor sells to the gray market (#1 and #4), total gray market sales decline. In this case, the unauthorized retailer is able to boost profits by forgoing some sales and waiting for a cheaper supply from the authorized retailer.

Table 7: Percent change in expected sales in the unauthorized channel

#	λ_u/λ_a	C_3/C_1	C_2/C_1	$(P_a - C_1)/P_a$					
				0.4	0.5	0.6	0.7	0.8	0.9
1	0.5	0.5	1.2	-36.68%	-9.58%	-4.02%	-2.13%	-1.13%	-0.34%
2	0.5	1.2	0.5	0.73%	0.69%	0.62%	0.55%	0.44%	0.41%
3	0.5	1.2	1.2	4.68%	2.16%	1.14%	0.53%	0.24%	0.19%
4	0.5	0.5	1.0	-3.52%	-2.09%	-1.37%	-0.86%	-0.46%	-0.06%
5	0.5	1.0	1.0	1.30%	0.83%	0.47%	0.29%	0.18%	0.18%
6	0.5	1.2	1.0	2.54%	1.61%	0.99%	0.61%	0.37%	0.27%

Impact on distributor sales of the authorized channel selling to the gray market. Here too, as shown in Table 8, we see that when the authorized channel sells at a price lower than that of the distributor, total distributor sales decline.

Table 8. Percent change in distributor sales

#	λ_u/λ_a	C_3/C_1	C_2/C_1	$(P_a - C_1)/P_a$					
				0.4	0.5	0.6	0.7	0.8	0.9
1	0.5	0.5	1.2	-8.51%	-1.91%	-0.78%	-0.47%	-0.28%	-0.09%
2	0.5	1.2	0.5	0.64%	0.50%	0.37%	0.27%	0.19%	0.15%
3	0.5	1.2	1.2	4.59%	2.23%	1.15%	0.57%	0.23%	0.10%
4	0.5	0.5	1.0	-0.14%	-0.11%	-0.14%	-0.12%	-0.09%	0.00%
5	0.5	1.0	1.0	1.97%	1.14%	0.64%	0.34%	0.17%	0.09%
6	0.5	1.2	1.0	2.55%	1.51%	0.85%	0.49%	0.24%	0.12%

Our results suggest that the unauthorized retailer will prefer to buy directly from the distributor unless retail margins are very low, but she will still choose to wait until demand uncertainty is resolved and then purchase from the authorized retailer. If the unauthorized channel can buy from the authorized retailer at a lower price than the distributor’s price, expected sales to the gray market decline relative to the level they would reach if the authorized retailer did not sell to the gray market. In this case, total distributor sales also decline.

These tables suggest complex dynamics among the three entities, resulting in flows that are not intuitive or easily explained. In general, the unauthorized retailer has to balance costs, supply risk, and demand risk. For the gray market, buying from the distributor increases the risk of making a commitment before obtaining demand information; on the other hand, waiting for the authorized retailer increases the supply uncertainty.

6. Concluding Remarks

We considered stochastic supply chain models in which an authorized retailer can divert unsold inventory to an unauthorized or gray market distribution channel. In these models, the ability to divert goods benefits the distributors and retailers. The practice can either benefit or harm producers, who will see additional demand but also will bear a number of potential costs and risks.

Specifically, we showed how the expected quantity diverted changes with expected authorized and unauthorized demand and with margins in various channels. In the process, we also

identified a number of counterintuitive relationships. For example, finding that even when penalties for sales to the gray market are large, the producer may still benefit.

Looking at two managerial leverages to mitigate the effects of gray markets, buyback, and multiple replenishment, we showed how buyback contracts could increase producer profit but prove unsatisfactory to authorized retailers. Multiple replenishments bring total quantity purchased closer to the true demand, though, benefitting the authorized retailer, and perhaps penalizing the producer.

When a distributor intermediates between the two types of retailers, we saw that the distributor's very presence reduces the ability of the producer to affect the gray market. We showed that, although the bulk of sales to the gray market are made by the distributor, sales to the gray market by the authorized retailer increase expected sales in the authorized channel. Here, too, the findings were counter-intuitive. For example, if the authorized retailer lowers his selling price to the gray market, expected gray market sales may decline. Building on this research, an interesting future task would be to study how pricing and product allocation can be used by producers to cope with gray markets.

***Acknowledgments:** We thank Raj Rajagopalan, Shantanu Dutta, Guillaume Roels, and three anonymous referees for their valuable comments on the previous version of our paper.*

References

Ahmadi, R., B. R. Yang. 2000. Parallel imports: Challenges from unauthorized distribution channels. *Marketing Science*, 19, 279-294.

Alliance for Gray Market and Counterfeit Abatement (AGMA). 2004. What is gray market or illegal brokering? AGMA. <http://www.agmaglobal.org/faq.html#graymarket1>.

Antia, K. D., S. Dutta, M. Bergen. 2004. Competing with gray markets. *Sloan Management Review*, 46(1), 63-69.

- Betancourt R. R. and D. A. Gautschi, 1993. The Outputs of Retail Activities: Concepts, Measurement and Evidence from U.S. Census Data, *The Review of Economics and Statistics* Vol. 75, No. 2, May, pp. 294-301.
- Cachon, G. P. 2003. Supply chain coordination with contracts. Supply chain management: Design, coordination and operation in Handbooks in Operations Research and Management Science, Vol. 11. S. C. Graves, A. G. De Kok (eds.). Elsevier Publishing Company, 229-340.
- Chen, H. 2009. Gray marketing: Does it hurt the manufacturers? *Atlantic Economic Journal*, 37, 23-35.
- Duhan, D.F., M.J. Sheffet. 1988. Gray markets and the legal status of parallel importation. *Journal of Marketing*, 52, 75-83.
- Dong, L., E. Durbin. 2005. Market for surplus components with a strategic supplier. *Naval Research Logistics*, 52, 734 – 753.
- Fisher, M. L., A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Operations Research*, 44, 87-99.
- Ganslandt, G., K. E. Maskus. 2004. Parallel imports and the pricing of pharmaceutical products: Evidence from the European Union. *Journal of Health Economics*, 23, 1035-1057.
- Gerstner, E., J. D. Hess, D. M. Holthausen. 1994. Price discrimination through a distribution channel: Theory and evidence. *The American Economic Review*, 84 (5), 1437-1445.
- Kouvelis, P, G. Gutierrez. 1997. The newsvendor problem in a global market: Optimal centralized and decentralized control policies for two-market stochastic inventory system, *Management Science*. 43(5), 571–585.
- Lee, H, S. Whang. 2002, The impact of the secondary market on the supply chain, *Management Science*, 48(6), 719-731.
- Maskus, K. E. 2000. Parallel imports. *World Economy*, (23), 1269-1284.
- Michael, J. 1998. A supplemental distribution channel?: The case of U.S. parallel export channels. *Multinational Business Review*, (6), 24-35.
- Milner, J. M., P. Kouvelis. 2007. Inventory, speculation, and sourcing strategies in the presence of online exchange. *MSOM*, 9(3), 312-331.
- Myers, M. B. 1999. Incidents of gray market activity among U.S. exporters: Occurrences, characteristics, and consequences. *Journal of International Business Studies*, 30, 105-126.

National Economic Research Associates (NERA). 1999. The economic consequences of the choice of regime in exhaustion in the area of trademarks.

Patel, J. K., C. Read. 1982. Handbook of the normal distribution. Marcel Dekker.

Rudi, N., S. Kapur and D. Pyke. 2001. A two-location inventory model with transshipment and local decision making, *Management Science*, 47, 1668 -1680.

Rudin, W. 1976. Principles of mathematical analysis (3rd ed.). McGraw Hill.

Kadiyali, V., P. Chintagunta and N. Vilcassim. 2000. Manufacturer-Retailer Channel Interactions and Implications for Channel Power: An Empirical Investigation of Pricing in a Local Market. *Marketing Science*, Vol. 19, No. 2, Spring, pp. 127-148.

Wee, K.E., M. Dada. 2005. Optimal policies for transshipping inventory in a retail network. *Management Science*, 51, 1519 – 1533.

Zhang, J., J Chen. 2003. Studies on the value of secondary market to the supply chain. *Journal of Systems Science and Systems Engineering*, 12(3), 360- 371.

Appendix: Data and Proofs

Demand distributions for Table 4: Authorized demand in the first period is $N(100,34.64)$ and second period is $N(100,30)$ with a correlation of 0.5. The unauthorized demand is $N(100,15)$ and the correlation between unauthorized and authorized markets in 0.4.

Proof of Proposition 1: Let, $m_a = P_a - C_1$, $m_u = C_3 - C_1$, and $m_s = v - C_1$. With the objective of maximizing expected margins aggregated across all three channels, the distributor's problem is to find the order quantity, I_a , that maximizes her aggregate expected margins, V , computed as:

$$V = E[m_a(I_a - I_u^+) + m_u(I_u^+ - I_s^+) + m_s I_s^+] \quad (1a) \text{ and rearranging, } = m_a I_a - (m_a - m_u)E[I_u^+] - (m_u - m_s)E[I_s^+].$$

$$(i): \text{ By the chain rule, } \frac{dE[I_u^+]}{d\lambda_a} = \frac{\partial E[I_u^+]}{\partial \lambda_a} + \frac{\partial E[I_u^+]}{\partial I_a} \cdot \frac{dI_a^*}{d\lambda_a}. \quad (2a).$$

$$\text{Taking partials of } E[I_u^+]: -\frac{\partial E[I_u^+]}{\partial \lambda_a} = \frac{\partial E[I_u^+]}{\partial I_a} = \Phi\left(\frac{I_a - \lambda_a}{\sigma_a}\right).$$

By the implicit function theorem $\frac{dI_a^*}{d\lambda_a} = \frac{-\frac{\partial^2 V}{\partial I_a \partial \lambda_a}}{\frac{\partial^2 V}{\partial I_a^2}}$. Differentiation of V gives:

$$\frac{\partial^2 V}{\partial I_a \partial \lambda_a} = -\frac{\partial^2 V}{\partial I_a^2} = \frac{m_a - m_u}{\sigma_a} \phi\left(\frac{I_a - \lambda_a}{\sigma_a}\right) + \frac{m_u - m_s}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}} \Phi\left(\frac{I_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right), \text{ so } \frac{dI_a^*}{d\lambda_a} \text{ equals 1. Thus, all}$$

terms on the right side of (2a) cancel; hence the constancy.

(ii) We prove similarly.

$$\frac{dE[I_u^+]}{d\lambda_u} = \frac{\partial E[I_u^+]}{\partial \lambda_u} + \frac{\partial E[I_u^+]}{\partial I_a} \cdot \frac{dI_a^*}{d\lambda_u} \text{ with } \frac{dI_a^*}{d\lambda_a} = \frac{-\frac{\partial^2 V}{\partial I_a \partial \lambda_u}}{\frac{\partial^2 V}{\partial I_a^2}} \quad (3a)$$

and we require that this is positive. Now, $\frac{\partial E[I_u^+]}{\partial \lambda_u}$ equals 0 by inspection of (2), $\frac{\partial E[I_u^+]}{\partial I_a}$ (given just above) is strictly positive as Φ is a cdf; and, $\frac{\partial^2 V}{\partial I_a^2}$ is negative by the concavity of V and (3a) is

$$\text{positive iff } \frac{\partial^2 V}{\partial I_a \partial \lambda_u}. \text{ Differentiating, } \frac{\partial^2 V}{\partial I_a \partial \lambda_u} = \frac{m_u - m_s}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}} \Phi\left(\frac{I_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right), \text{ which is indeed}$$

positive. Proof of (iii) is omitted, as it is nearly identical to this. QED

Proof of Proposition 2: (i) We require that $E[I_u^+]$ increases with σ_a if and only if $I_a^* > \lambda_a + \lambda_u$.

The chain rule and the implicit function theorem (applied to I_a^*) gives $\frac{dE[I_u^+]}{d\sigma_u} = \frac{\partial E[I_u^+]}{\partial \sigma_u} + \frac{\partial E[I_u^+]}{\partial I_a}$.

$\frac{dI_a^*}{d\sigma_u}$ with $\frac{dI_a^*}{d\sigma_a} = \frac{-\frac{\partial^2 V}{\partial I_a \partial \sigma_u}}{\frac{\partial^2 V}{\partial I_a^2}}$. By inspection of (2), $\frac{\partial E[I_u^+]}{\partial \sigma_u} = 0$; $\frac{\partial E[I_u^+]}{\partial I_a} = \Phi\left(\frac{I_a - \lambda_a}{\sigma_a}\right)$ which is always positive,

and $\frac{\partial^2 V}{\partial I_a^2}$ is negative by the concavity of V in I_a . So, $E[I_u^+]$ increases with σ_u if and only if $\frac{\partial^2 V}{\partial I_a \partial \sigma_u}$ is

$$\text{positive. } \frac{\partial^2 V}{\partial I_a \partial \sigma_u} = (m_u - m_s) \frac{\sigma_u + \rho\sigma_a}{(\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u)^{\frac{3}{2}}} (I_a - \lambda_a - \lambda_u) \phi\left(\frac{I_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right), (m_u - m_s) \text{ is}$$

positive by assumption; $\phi(\cdot)$ is positive as it is a pdf; and σ_a and σ_u are standard deviations. So, the expression is positive if and only if $I_a > \lambda_a + \lambda_u$ as required.

(ii) We need to prove that $\frac{dE[I_u^+]}{d\sigma_u} > 0$:

$$\frac{dE[I_u^+]}{d\sigma_a} = \frac{\partial E[I_u^+]}{\partial \sigma_a} + \frac{\partial E[I_u^+]}{\partial I_a} \cdot \frac{dI_a^*}{d\sigma_a} \text{ with } \frac{dI_a^*}{d\sigma_a} = \frac{-\frac{\partial^2 V}{\partial I_a \partial \sigma_a}}{\frac{\partial^2 V}{\partial I_a^2}}. \quad (4a)$$

In this case, $\frac{\partial E[I_u^+]}{\partial \sigma_a}$ equals $\phi\left(\frac{I_a - \lambda_a}{\sigma_a}\right)$ and is positive; $\frac{\partial E[I_u^+]}{\partial I_a}$ equals $\Phi\left(\frac{I_a - \lambda_a}{\sigma_a}\right)$ and is also positive; $\frac{\partial^2 V}{\partial I_a^2}$ is negative by concavity; and

$$\frac{\partial^2 V}{\partial I_a \partial \sigma_a} = (m_a - m_u) \frac{1}{\sigma_a^2} (I_a - \lambda_a) \phi\left(\frac{I_a - \lambda_a}{\sigma_a}\right) + (m_u - m_s) \frac{\sigma_u + \rho\sigma_a}{(\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u)^{\frac{3}{2}}} (I_a - \lambda_a - \lambda_u) \phi\left(\frac{I_a - \lambda_a - \lambda_u}{\sqrt{\sigma_a^2 + \sigma_u^2 + 2\rho\sigma_a\sigma_u}}\right)$$

$I_a - \lambda_a > I_a - \lambda_a - \lambda_u$ together with $(m_a - m_u)$, $(m_u - m_s)$, $\phi(\cdot)$, σ_a , and σ_u are all positive, which implies that this expression and (4a) are positive (as required) whenever $I_a > \lambda_a + \lambda_u$. QED

Proof of Proposition 3: The producer's expected profits are differentiable, and Q_a is chosen from

an interval, so it suffices to find conditions for which $\frac{\partial}{\partial Q_a} B_p > 0$. $\frac{\partial}{\partial I_a} B_p = M - \pi \cdot \left\{ \frac{\partial}{\partial Q_a} E[I_u^+] -$

$$\frac{\partial}{\partial Q_a} E[I_s^+] \right\} \frac{\partial}{\partial Q_a} E[I_u^+] = \frac{\partial}{\partial Q_a} \int_{x=0}^{Q_a} (Q_a - x) f_a(x) dx = F_a(Q_a) \frac{\partial}{\partial Q_a} E[I_s^+] = \frac{\partial}{\partial Q_a} \int_{x=0}^{Q_a} (Q_a -$$

$x) f_{a+u}(x) dx = F_{a+u}(Q_a)$. Therefore: $\frac{\partial B_p}{\partial Q_a} > 0$ implies $\frac{M}{\pi} > F_a(Q_a) - F_{a+u}(Q_a)$. QED

Proof of Proposition 4: Equation (6) corresponds to the optimality condition for the second

purchase decision for the authorized retailer and $Q_{2,a}^*$ is the order-up-to level for the second

purchase. The authorized retailer's earnings in the second half are a convex function of the starting

inventory. If the starting inventory at the beginning of the second half is Q , demand for the second

half is d_1 , gray market demand is d_2 , then the authorized retailer's earnings are:

$$R_2(Q, d_1, d_2) = P_a \text{Min}(Q, d_1) + C_3 \text{Min}\{\text{Max}[0, (Q - d_1)], d_2\} + v \text{Max}[0, (Q - d_1 - d_2)].$$

Because $P_a > C_3 > v$, $R(Q, d_1, d_2)$ is concave in Q and the expected earnings for the second half

are a concave function of Q . The first order conditions given by (6) are sufficient and the optimal

second purchase quantity is given by Equation (7).

If the initial purchase is Q and demand in the first half is d , then the expected earnings are:

$$R_1(Q, d) = P_a \text{Min}(Q, d) + \text{Max}_{x \geq 0} E_{d_1, d_2} R_2(\text{Max}[0, Q - d] + x, d_1, d_2).$$

Once again, $R_1(\cdot; \cdot)$ is a concave function of d and the first order conditions in Equation (8) are sufficient. Equation (8) provides the expected marginal value of additional purchases at the beginning of the season, given that $Q_{1,a}$ units have already been purchased by the authorized retailer. The first term corresponds to the probability that the demand in the first half exceeds $Q_{1,a}$. Any additional purchases would fetch the selling price, P_a .

The remaining terms in (8) correspond to situations in which there is some inventory available at the end of the first half. If the inventory is less than the order-up-to level, then a second purchase brings the starting inventory back up to the order-up-to level. The incremental value in this case is C_1 . If the starting inventory exceeds the order-up-to level, then no purchases are made at the end of the first half. The incremental value will then depend on whether the leftover inventory is enough to meet the demand in the second half. The last three terms correspond to situations when the incremental unit purchased in the first half is used to service demand in the (i) authorized market during the second half, (ii) gray market, and (iii) salvage market, respectively.

All that is left to show is that if demand is less than x^* given by Equation (9), then the residual inventory exceeds the order-up-to level for the second half. Demand in the two halves is bivariate normal with positive correlation. The mean of the demand in the second half is monotone increasing in the demand in the first half. As a result, $Q_{2,a}(x)$ is monotone increasing in x .

Consequently, for all $x \leq x^*$, $Q_{1,a} - x \geq Q_{2,a}(x)$. QED

Proof of Proposition 5: Consider the profit function conditioned on D_a and D_u :

$\Pi_a(Q_a, Q_u | D_a, D_u)$. Fixing D_u is equivalent to fixing Y . If $Q_a < D_a$, then the incremental revenue is P_a . Once Q_a exceeds D_a , **a** either salvages or sells at C_3 . If he sells at C_3 , there is another threshold, T^* , that only depends on D_a and D_u , such that if Q_a exceeds T^* , **a** will salvage. Thus for given D_a

and D_u , $\Pi_a(\cdot, Q_u|D_a, D_u)$ is concave. The function $\Pi_a(Q_a, Q_u)$ is concave because convex combinations of concave functions are concave. Formally:

$$\frac{\partial \Pi_a(Q_a, Q_u)}{\partial Q_R} = P_a \int_{r=Q_a}^{\infty} f_a(r) dr + C_3 \int_{r=0}^{Q_R} \int_{y=Q_a-r}^{\infty} f_y(y|r, Q_a) f_a(r) dy dr + v \int_{r=0}^{Q_a} \int_{y=0}^{Q_a-r} f_y(y|r, Q_u) f_a(r) dy dr + v \int_{r=0}^{Q_R} Pr_Y(0|r, Q_u) f_a(r) dr - C_1$$

and

$$\frac{\partial^2 \Pi_a(Q_a, Q_u)}{\partial Q_a^2} = -(P_a - v Pr_Y(0|Q_a, Q_u) f_a(Q_a) - (C_3 - v) \int_{r=0}^{Q_R} f_y(Q_a - r, Q_u) f_a(r) dr < 0.$$

The strict inequality is due to Assumption A1. Consequently, $\Pi_a(Q_a, Q_u)$ is strictly concave in Q_a for fixed Q_u and the results follow. QED

Proof of Proposition 6: Once again, fix D_a and D_u and determine the incremental change in objective value as we increase Q_u . $Z = \text{Max}(0, Q_a - D_a)$. We have to consider two cases: **Case 1:** $Z \geq (1 - \theta)D_u$ and **Case 2:** $Z \leq (1 - \theta)D_u$.

When $Q_u = 0$, no sales will occur during Period 1. Product is bought at price C_3 to cover demand in Period 2. The quantity purchased will depend on the availability, which is known once we fix Q_a and D_a . As we increase Q_u , the unauthorized retailer sells at P_u . Once she covers demand during Period 1, she will either sell more in Period 2 or reduce the quantity bought from \mathbf{a} . Finally, she will salvage the remaining product. Thus, the incremental value is non-increasing. Formally:

$$\begin{aligned} \frac{\partial \Pi_u(Q_a, Q_u)}{\partial Q_u} &= P_u \int_{s=\frac{Q_u}{\alpha}}^{\infty} f_u(s) ds + P_u \int_{s=Q_u}^{\frac{Q_u}{\alpha}} \left(Pr_Z(0|s, Q_a) + \int_{z=0}^{s-Q_u} f_z(z|s, Q_a) \right) f_u(s) dz ds + C_3 \int_{s=Q_u}^{\frac{Q_u}{\alpha}} \int_{z=s-Q_u}^{\infty} f_z(z|s, Q_u) f_u(s) dz ds + v \int_{s=0}^{Q_s} f_u(s) ds - C_2 \\ \frac{\partial^2 \Pi_u(Q_a, Q_u)}{\partial Q_u^2} &= -\frac{1}{\alpha} [P_u - C_3] f_u\left(\frac{Q_u}{\alpha}\right) \int_{r=0}^{Q_a - \frac{1-\alpha}{\alpha} Q_u} f_a\left(r|s = \frac{Q_u}{\alpha}\right) dr - (P_u - C_3) \int_{s=Q_u}^{\frac{Q_u}{\alpha}} f_a(Q_a + Q_u - s|s) F_u(s) ds - \left(C_3 \int_{r=0}^{Q_a} f_R(r|s = Q_u) dr + P_u \int_{r=Q_a}^{\infty} f_a(r|s = Q_u) dr - V \right) f_u(Q_u) < 0. \end{aligned} \quad (9a)$$

Once again, the strict inequality is due to Assumption A1. QED

Proof of Proposition 7: First consider $G_u(\cdot)$. We will show that if Q_a increases, $G_u(\cdot)$ will not increase. Z , the surplus available for sale at the end of Period 1, is non-decreasing in Q_a . Let us fix $Q_u = G_u(Q_a)$ and increase Q_a to $Q_a + \delta$. We want to show that the expected marginal value of Q_u

for retailer \mathbf{u} does not increase as Q_a increases. In all the scenarios in which no purchases are made from \mathbf{a} , potential change in Z will have no impact on the marginal value of Q_u for \mathbf{u} . Now consider scenarios in which there is a purchase from \mathbf{a} . Here, if Z was sufficient to meet demand, then an increase in Z will have no impact on the marginal value of Q_u . Then, \mathbf{u} has purchased all that she needs at the beginning of Period 2; therefore, any additional supply will have zero incremental value. Now, consider a situation in which Z is not adequate. In this scenario, an additional unit of Q_u would have been worth $(P_u - C_2)$. At the margin, an increase in Z will decrease the value of an additional unit of Q_u to $(C_3 - C_2)$. Therefore, the expected incremental value of Q_u does not increase in any scenario if Q_a increases. This in turn implies that the overall expected marginal value cannot increase and the purchase quantity cannot increase. Thus,

$$\frac{\partial^2 \Pi_u(Q_a, Q_u)}{\partial Q_u \partial Q_a} = -(P_u - C_3) \int_{s=Q_u}^{\frac{Q_u}{\theta}} f_a(r = Q_a + Q_u - s) f_u(s) ds. \quad (10a)$$

Thus $G_u(\cdot)$ is monotone non-increasing.

Similarly, an increase in Q_u will decrease the value of an additional unit of Q_a . The retailer, \mathbf{u} is less likely to buy the incremental unit. Hence $G_a(\cdot)$ is also monotone non-increasing.

For $Q_a \leq (1 - \theta)Q_u/\theta$: $\frac{\partial^2 \Pi_a(Q_u, Q_a)}{\partial Q_a \partial Q_u} = -(C_3 - v) \int_{r=0}^{Q_a} f_u(s = Q_a + Q_u - r|r) f_a(r) dr$ (11a). For

$Q_a \geq (1 - \theta)Q_u/\theta$: $\frac{\partial^2 \Pi_a(Q_u, Q_a)}{\partial Q_a \partial Q_u} = -(C_3 - v) \int_{s=\frac{1-\theta}{\theta}Q_u}^{\frac{Q_u}{\theta}} f_u(s|r) f_a(r = Q_a - \frac{1-\theta}{\theta}Q_u) Q_u ds$. QED (12a)

$G_u(Q_a)$ and $G_a(Q_u)$ are both monotone non-increasing functions. If Q_a^* and Q_u^* are such that $G_u(Q_a^*) = G_a^{-1}(Q_u^*)$, then Q_a^* and Q_u^* represent equilibrium purchase quantities.

Proof of Proposition 8: If \mathbf{u} is assured of supply at price C_3 at the end of Period 1, she will purchase a quantity Q_u^∞ that maximizes the following:

$$\int_{s=\frac{Q_u}{\theta}}^{\infty} \{P_u Q_u + (1 - \theta)s\} f_u(s) ds + \int_{s=Q_u}^{\frac{Q_u}{\theta}} (P_u s - C_3(s - Q_u)) f_u(s) ds + \int_{s=0}^{Q_u} \{P_u s + v(Q_u - s)\} f_u(s) ds - C_2 Q_u. \quad (12a)$$

We know $G_u(\cdot)$ is monotone, therefore $\text{Lim}_{Q_a \rightarrow \infty} G_u(\cdot) \rightarrow Q_u^\infty > 0$.

On the other hand, if $Q_u = 0$, then \mathbf{u} will solve a newsvendor problem based on parameters P_u , C_2 , and v . Let Q_u^0 denote the newsvendor solution for \mathbf{u} ; therefore, $G_u(0) = Q_u^0 < \infty$.

If \mathbf{a} cannot sell to \mathbf{u} , then he too will solve a newsvendor problem with parameters, P_a , C_1 , and v .

Let Q_a^∞ denote the newsvendor solution for \mathbf{a} . $\text{Lim}_{Q_u \rightarrow \infty} G_a(\cdot) \rightarrow Q_a^\infty > 0$.

Finally, if $Q_u = 0$, then \mathbf{a} knows he can sell $(1 - \theta)s$ units to \mathbf{u} at the end of Period 1. Then, \mathbf{a} will purchase quantity Q_a^0 that will maximize the function given by (2), but now y represents the entire demand in Period 2 for \mathbf{u} ; therefore, $G_a(0) = Q_a^0 < \infty$.

Consider the graphs $(G_u(Q_a), Q_a)$ and $(Q_u, G_a(Q_u))$. These graphs must intersect because (i) $G_u(\cdot)$ and $G_a(\cdot)$ are continuous and (ii) $Q_u^\infty < G_u(Q_a) < Q_u^0 < \infty$ and $Q_a^\infty < G_a(Q_u) < Q_a^0 < \infty$.

Each intersection point is an equilibrium point as shown in Figure 4. By the implicit function theorem (Rudin 1976), functions $G_u(\cdot)$ and $G_a(\cdot)$ are locally continuous in neighborhoods in which

$\frac{\partial^2 \Pi_u(Q_a, Q_u)}{\partial Q_u \partial Q_a}$ and $\frac{\partial^2 \Pi_a(Q_u, Q_a)}{\partial Q_a \partial Q_u}$ are not equal to zero, respectively. QED

Figure 4. Graphs of $G_u(\cdot)$ and $G_a(\cdot)$

