WHAT CAN THE DURATION OF DISCOVERED CARTELS TELL US ABOUT THE DURATION OF ALL CARTELS?*

Joseph E. Harrington Jr. and Yanhao Wei

Estimates of average cartel duration and the annual probability of cartel death are based on data for discovered cartels. It is recognised that these estimates could be biased because the population of discovered cartels may not be a representative sample of the latent population of cartels. This article constructs a birth-death-discovery process to investigate the source and direction of possible biases. Bayesian inference is used to provide bounds on the extent of the bias and deliver an improved set of beliefs on the probability of cartel death.

It is well understood that cartels are socially harmful. What is far less clear is the magnitude of this harm. The seriousness of the cartel problem in an economy depends on:

(*i*) how many cartels there are;

(*ii*) how large is the overcharge (and the elasticity of market demand); and

(*iii*) how long cartels last.

Currently, what we know about these issues comes from cartels which, to their disappointment, were discovered and convicted. Other than that, the number of discovered cartels is a lower bound on the number of cartels, we know little about how many cartels there are.[1] On the issue of the overcharge, there is a fair amount of work estimating how much higher is the price charged by discovered cartels. For surveys of those estimates, see Connor and Bolotova (2006), Oxera (2009), Connor (2010) and Boyer and Kotchoni (2012). Regarding the last factor – how long cartels last – the consensus measure of average cartel duration is the average duration of discovered cartels which most studies find to be five to seven years.[2]

Cartel duration is not only relevant to assessing the severity of the cartel problem but also to evaluating whether current levels of enforcement are sufficient to deter cartel formation. Some economists have expressed a lack of concern about cartels on the premise that cartels are inherently unstable and thus, even if they form, they will not be around for long. That is, of course, an empirical question. There is also an active

[1] While this does not pertain to the question of how many cartels there are now, there is research that addresses how many cartels there were when cartels were lawful. At various points in time, Austria, Denmark, Germany, Finland, Norway and Sweden permitted cartels but required that they be registered with the government. For an analysis of the Finnish cartel registry, see Hyytinen *et al.* (2014*a*, *b*).

[2] Levenstein and Suslow (2006) summarise the findings of many studies on cartel duration.

debate regarding whether penalties in some jurisdictions are too low – so that there is under-deterrence of cartels – or too high – so that they are in excess of what is necessary to deter and may be creating social costs. For example, Connor and Lande (2012) argues that there is under-deterrence. In the case of the European Union, Allain *et al* (2011) provides evidence against the under-deterrence claim, though see Combe and Monnier (2011) for a different view.[3] An evaluation of the extent to which enforcement is either under or over-deterring cartels depends on the incremental profits earned from collusion, penalties in the event of discovery and conviction, and the likelihood of discovery and conviction. Using data on the duration of discovered cartels, these and other studies use an estimate of 15% per year that a cartel is caught and convicted.[4]

In assessing the extent of the cartel problem and determining whether enforcement should be strengthened or weakened, estimates of the average duration of discovered cartels and the annual probability of discovery and conviction for discovered cartels are useful only to the extent that they are reasonable proxies for the average duration of all cartels (discovered and undiscovered) and the annual probability of discovery and conviction for all cartels (discovered and undiscovered). The objective of this article is to explore to what extent they are reasonable proxies. Are estimates using data on discovered cartels biased? If so, what is the direction and possible magnitude of the bias?

Section 1 describes the stochastic process governing cartel birth, death and discovery. The theoretical construct is explored in Sections 2 and 3 to provide insight into the sources of bias and how these sources affect the direction of bias. Contrary to common perception (for which we later provide references), we show that average duration of discovered cartels need not be a biased measure of average cartel duration. In fact, if the empirical model used in previous studies is correctly specified then estimates are unbiased. That result is fairly intuitive and not particularly surprising. Of greater value is showing how, under plausible assumptions, estimates based on data using discovered cartels are likely to be biased and then characterising when average duration of discovered cartels is an overestimate and when it is an underestimate of average cartel duration. In Section 4, this theoretical model is used along with data on discovered cartels to provide bounds on the extent of bias through the application of Bayesian inference. While there is evidence of bias, it proves to be modest in magnitude. Finally, we deliver an improved set of beliefs on the probability of cartel death that takes into account this bias. A cartel has, on average, a 17% chance each year of either collapsing and/or being discovered.

---

[3] Also see Harrington (2014*a*) who argues that enforcement policy is more effective than is generally believed. For some reasons as to why we should not ease up on enforcement even if we think penalties are in the over-deterrence region, see Harrington (2014*b*).

[4] For 184 convictions by the Antitrust Division of the US Department of Justice over 1961–88, Bryant and Eckard (1991) estimated the annual chances of disovery and conviction to lie between 13% and 17%, while for 86 convictions by the European Commission over 1969–2007, Combe *et al.* (2008) estimated it to be around 13%. For more than 200 cartels from more than 20 developing countries over 1995–2013, Ivaldi *et al.* (2015) estimate a higher rate of 24%.

While the analysis is motivated by learning about cartels, the ensuing theoretical and empirical analyses could be applicable to other forms of illegal conspiracy such as drug gangs and counterfeiting rings. As long as they are subject to stochastic death and discovery, the observed population of illegal conspiracies may be a biased sample of all such conspiracies. The results derived in this article could then shed light on what one can infer from detected unlawful conspiracies about those that avoid detection.

1. Model

There are generally recognised to be two primary reasons why the duration of discovered cartels may be a biased measure of the duration of all cartels: measurement error and sample selection bias. Cartel duration is measured as the time between a cartel's birth and its death. Cartel death is typically well documented, though there can be cases in which cartels remain active even after discovery.[5] Far more problematic is dating the birth of a cartel. In most data sets, a cartel's date of birth is either the earliest time for which there is evidence of a cartel or is the product of negotiation between the defendants and the competition authority. In either case, it is reasonable to presume that official cartel birth is no earlier than actual cartel birth and, therefore, any bias from measurement error is likely to result in the duration of discovered cartels' being an underestimate.

In this article, we do not address the issue of measurement error and instead presume that the duration of discovered cartels is accurately measured. Our focus is on characterising selection bias associated with using the population of discovered cartels to draw inferences about the population of cartels. In order to produce some clean insight, the noise associated with finite samples is also not considered, as it is assumed there is an infinite number of industries of which some fraction are cartelised out of which some fraction are discovered. We are then characterising what is referred to as asymptotic bias, which is the bias present with infinite samples.[6]

For a continuum of industries, assume there is a cartel birth process that is constant over time so that, in each period, a mass of cartels are created, which is normalised to one.[7] Once cartelised, a cartel can 'die' for natural reasons (due to a change in market conditions or firm-specific factors so collusion is no longer stable) or it can die because it has been detected and convicted by the competition authority (which, for brevity, will hereon be referred to as 'discovery'). A cartel can then have three possible terminal states:

---

[5] There are some well-documented cases of post-conviction collusion in government procurement auctions in Japan (Kawai and Nakabayashi, 2014).

[6] An assumption made throughout the analysis is that we can only observe whether a cartel was discovered but not whether, upon discovery, it was active or inactive. While that assumption fits most empirical studies, such information is available in some instances (De, 2010; Levenstein and Suslow, 2011). However, measurement issue is a concern. For example, suppose a customer complained and firms learned about it which then caused the cartel to collapse. Thus, at the time a formal investigation is opened up, the cartel would be inactive and one might be inclined to conclude that the cartel collapsed and was then discovered but the reality is that discovery – in the form of a customer complaint, not the formal investigation – caused the collapse and, therefore, the cartel was actually active at the time of discovery.

[7] Allowing the birth process to be time-dependent (perhaps sensitive to the business cycle) is a worthy extension but, at this stage, it would be counter-productive to gaining some initial insight into selection bias.

(*i*) it can die a natural death (which we refer to as 'collapse') and not be discovered;

(*ii*) it can collapse and be discovered; and

(*iii*) it can be discovered (and thus die through conviction).

The probability that a cartel collapses in the current period is λ, the probability of discovery conditional on the cartel being active in the current period is ρ, and the probability of discovery conditional on the cartel just having collapsed in the current period is β. There are a variety of reasons why the likelihood of discovery may depend on whether the cartel is still active or has collapsed. For example, if customer complaints are triggered by suspicious price movements then a sharp price decline associated with the cartel collapsing may trigger discovery.[8] However, the primary reason why we have allowed for this possible dependence of discovery on the state of the cartel is to take account of leniency programmes. A deterrent to a member of an active cartel applying for leniency is that it will cause the cartel to collapse and, as a result, the firm will forego future collusive profits. However, if the cartel is no longer active then that cost disappears in which case firms may then self-report.[9] This would provide a rationale for $\beta > \rho$ so that cartel collapse makes discovery more likely. Alternatively, it could be the case that $\beta = \rho$ so the probability of discovery is the same whether the cartel is alive or just died. At this stage, it is assumed: $\lambda \in (0,1)$, $\rho \in (0,1)$, $\beta \in [0,1]$.

Studies that estimate the annual probability of discovery for discovered cartels – such as Bryant and Eckard (1991) and Combe *et al.* (2008) – specify a constant hazard rate model. This is a continuous time model for which there is a constant probability of being caught in any instant. The model of this article is a discrete time analogue in that the probabilities of death and discovery are also constant over time. In addition, it is worth noting that results can be stated either in terms of average cartel duration or the probability of death in that the former equals the inverse of the latter. For example, an over-estimate of average cartel duration is equivalent to an underestimate of the probability of death.

Before moving on, let us note that the possibility of selection bias associated with using data on discovered cartels is recognised, though different views have been expressed regarding the direction of bias. In the original study of Bryant and Eckard (1991, pp. 535–6), it was suggested (though not presumed) that 'the life of a caught conspiracy is typically no longer than that of an uncaught conspiracy ... and [so] our probability [of death] estimate is an upper bound'. This suggestion that the probability of death is over-estimated was picked up more recently in Connor and Lande (2012, pp. 462–3): '[a cartel's probability of death] p is computed from samples of discovered cartels. Founders of never-discovered cartels might rationally conjecture a lower p. Thus, computed sizes of p may well overstate the actual average p for all cartels'. This direction to the bias is based on the presumption that if a cartel was not caught then it would have survived longer and,

---

[8] The implications for cartel pricing of having discovery depend on past prices is explored in Harrington (2004, 2005).

[9] The birth and death of cartels is modelled in Harrington and Chang (2015) and, under the assumption of full leniency, all dying cartels have firms racing for leniency which, for the current model, means $\beta = 1$.

therefore, the duration of discovered cartels is less than that for undiscovered cartels. While those papers emphasise that cartel duration is underestimated, Levenstein and Suslow (2011, p. 463) believe that cartel duration is overestimated: 'Because our sampling procedure relies on antitrust prosecutions, it is less likely to capture very short lived cartels. These may form and disappear without ever attracting the attention of the authorities. Thus, as with most other samples of cartels, our estimates of cartel duration may be biased upward relative to the universe of all cartels ever attempted'.

In light of the disparity of views, the initial motivation for this project was to examine the matter rigorously in order to inject some clarity. In Section 2, it is assumed that all cartels are produced by the same process; that is, all cartels have the same (λ, ρ, β). Thus, cartels are *ex ante* the same but are *ex post* different because they have different realisations of the death-discovery stochastic process. This is essentially the model that underlies studies like Bryant and Eckard (1991). We find, contrary to the various views just expressed, there is no bias: average cartel duration for discovered cartels equals average cartel duration. However, inferences drawn by previous studies regarding the probability of discovery are mistaken. In Section 3, the more reasonable assumption is made which is that cartels are governed by different processes; that is, cartels can have different values for (λ, ρ, β). Now there is bias and we investigate the determinants of the direction and size of the bias.

2. Cartels are Subject to the Same Death and Discovery Process

Suppose a mass 1 of cartels are born each period and that has been true for the infinite past; hence, the cartel population is in the steady-state. There is then a mass of cartels that are born in the current period out of which $(1 - \lambda)\rho + \lambda$ die – a fraction λ collapse and a fraction $(1 - \lambda)\rho$ do not collapse but are discovered and thus die – and $(1 - \lambda)\rho + \lambda\beta$ are discovered – a fraction $\lambda\beta$ collapse and are discovered and a fraction $(1 - \lambda)\rho$ do not collapse and are discovered. Given that a cartel that died with duration of one period must have experienced birth and death in the current period (born at the start and died at the end) then the mass of cartels with duration 1 is $(1 - \lambda)\rho + \lambda$. Analogously, the mass of cartels discovered with duration 1 is $(1 - \lambda)\rho + \lambda\beta$.

Cartels with a duration of two periods must have been born one period ago, survived for that period, and then died in the current period. With a mass 1 of cartels born a period ago, a mass $(1 - \lambda)(1 - \rho)$ of them survived to the current period – they neither collapsed nor were discovered – and, of those, a fraction $(1 - \lambda)\rho + \lambda$ die this period which implies there is a mass $[(1 - \lambda)\rho + \lambda](1 - \lambda)(1 - \rho)$ of cartels that lasted two periods and, analogously, a mass $[(1 - \lambda)\rho + \lambda\beta](1 - \lambda)(1 - \rho)$ of cartels that are discovered after two periods.

These values are shown in Table 1 along with values for other durations. Given that all cartels eventually die, average cartel duration is just the sum of the numbers in the column 'Mass of cartels of duration t that died in the current period' with each number weighted by the length of cartel duration:

Table 1

Population of Cartels in the Current Period

t	Mass of cartels born t periods ago and are alive in the current period	Mass of cartels of duration t that died in the current period	Mass of cartels of duration t discovered in the current period
1	1	$(1 - \lambda)\rho + \lambda$	$(1 - \lambda)\rho + \lambda\beta$
2	$(1 - \lambda)(1 - \rho)$	$(1 - \lambda)(1 - \rho)[(1 - \lambda)\rho + \lambda]$	$(1 - \lambda)(1 - \rho)[(1 - \lambda)\rho + \lambda\beta]$
3	$[(1 - \lambda)(1 - \rho)]^2$	$[(1 - \lambda)(1 - \rho)]^2[(1 - \lambda)\rho + \lambda]$	$[(1 - \lambda)(1 - \rho)]^2[(1 - \lambda)\rho + \lambda\beta]$
⋮			
T	$[(1 - \lambda)(1 - \rho)]^{T-1}$	$[(1 - \lambda)(1 - \rho)]^{T-1}[(1 - \lambda)\rho + \lambda]$	$[(1 - \lambda)(1 - \rho)]^{T-1}[(1 - \lambda)\rho + \lambda\beta]$
Sum		$[(1 - \lambda)\rho + \lambda]\sum_{t=1}^{\infty}$ $[(1 - \lambda)(1 - \rho)]^{t-1} = 1$	$[(1 - \lambda)\rho + \lambda\beta]\sum_{t=1}^{\infty}[(1 - \lambda)(1 - \rho)]^{t-1}$ $= \dfrac{(1 - \lambda)\rho + \lambda\beta}{1 - (1 - \lambda)(1 - \rho)}$

$$\sum_{t=1}^{\infty} t[(1 - \lambda)\rho + \lambda][(1 - \lambda)(1 - \rho)]^{t-1} = [(1 - \lambda)\rho + \lambda]\sum_{t=1}^{\infty} t[(1 - \lambda)(1 - \rho)]^{t-1}$$

$$= [(1 - \lambda)\rho + \lambda]\left\{\frac{1}{[(1 - \lambda)\rho + \lambda]^2}\right\}$$

$$= \frac{1}{\lambda + \rho - \lambda\rho}.$$

Average cartel duration is then $(\lambda + \rho - \lambda\rho)^{-1}$.

Next, let us turn to the average duration of discovered cartels which we will also refer to as average discovery time (conditional on being discovered). Recognising that only a mass $[(1 - \lambda)\rho + \lambda\beta]/[1 - (1 - \lambda)(1 - \rho)]$ of all cartels are discovered, average time until discovery is the sum of the numbers in the column 'Mass of cartels of duration t discovered in the current period' divided by the total mass of discovered cartels with each weighted by the length of duration:[10]

$$\sum_{t=1}^{\infty} t\left\{\frac{[(1 - \lambda)\rho + \lambda\beta][(1 - \lambda)(1 - \rho)]^{t-1}}{\dfrac{(1 - \lambda)\rho + \lambda\beta}{1 - (1 - \lambda)(1 - \rho)}}\right\} = \sum_{t=1}^{\infty} t\left\{\frac{[(1 - \lambda)(1 - \rho)]^{t-1}}{\dfrac{1}{1 - (1 - \lambda)(1 - \rho)}}\right\}$$

$$= \sum_{t=1}^{\infty} t[(1 - \lambda)\rho + \lambda][(1 - \lambda)(1 - \rho)]^{t-1} = \frac{1}{1 - (1 - \lambda)(1 - \rho)} = \frac{1}{\lambda + \rho - \lambda\rho}.$$

Average discovery time then equals average cartel duration.

PROPERTY 1. *If all cartels have the same probabilities of collapse and discovery, then average duration for discovered cartels (average discovery time) equals average duration for all cartels.*

This property may seem counter-intuitive because those cartels that are discovered would have continued to survive if they had not been discovered in which case one

[10] This calculation does presume $\rho > 0$ and/or $\beta > 0$ so that there is a positive mass of cartels that are discovered.

might think that the average time until death would necessarily exceed the average time until discovery. But the issue is not whether discovery shortens cartel life – it does – but rather whether discovery delivers a representative sample of the population of cartels. Inspecting Table 1, note that the discovery process samples a fraction $(1 - \lambda)\rho + \lambda\beta$ of all surviving cartels and, therefore, the population of discovered cartels is, indeed, a representative sample of the population of all cartels. Hence, the average age of cartels in the discovered population is the same as the average age of cartels in the entire population. Given that much of the article will involve loosening the restrictive assumptions that underlie that result, the takeaway should not be that average duration for discovered cartels is a good proxy for average cartel duration. Rather the takeaway is that the common perception that average duration for discovered cartels is necessarily a biased estimate of average cartel duration is incorrect.

Though, under the assumption of a common birth-death-discovery process, we can derive an unbiased measure of cartel duration, it is not possible to measure the likelihood that a cartel is discovered. What we can measure is the cartel survival rate $(1 - \lambda)(1 - \rho)$; that is, the probability that an active cartel neither collapses nor is discovered. Inspecting Table 1, the survival rate equals the ratio of the mass of discovered cartels in period $t + 1$ to the mass of discovered cartels in period t. Though the survival rate can be derived, the probability of collapse λ and the probability of discovery (conditional on being active) ρ cannot be separately identified, and nothing can be said about β (which is the probability of discovery conditional on having just collapsed).

The inability to disentangle λ and ρ is a point not well-appreciated. Assuming that the empirical model is correctly specified, the estimate of around 0.15 from studies such as Bryant and Eckard (1991) is an estimate of the probability of death, $1 - (1 - \lambda)(1 - \rho)$, not the probability of discovery and conviction, ρ. If there is no model misspecification, it is then an upper bound on ρ (which is only achieved when $\lambda = 0$ so cartels are perfectly stable). Yet, many studies appeal to this empirical work to justify assuming there is a 15% chance each year that a cartel is caught and convicted. For example, in showing that the socially optimal penalty is approximately 75% of the current formula used by the European Commission, Katsoulacos and Ulph (2013) assume the annual probability a cartel pays penalties is 0.15. Hinloopen and Soetevent (2008) engage in experimental work to investigate the impact of leniency programmes and assume that subjects who choose to communicate incur a penalty with probability 0.15. In discussing the role of customer damages, Sweeney (2006, p. 843) appeals to Bryant and Eckard (1991) to note that 'only about 13–17% of cartels in the US are detected'. And, until deriving this result, one of the authors of this article was equally guilty of this error (Harrington, 2014a).

Another misconception is that these estimates allow one to infer how many cartels escape discovery. As a recent example, an *Amicus Curiae* Brief submitted to the US Seventh Circuit Court refers to 'estimates suggesting that more than two-thirds of conspiratorial activity goes undetected and unpunished'.[11] If we use the expression in

---

[11] '*Amicus Curiae* Brief of Economists and Professors in Support of Appellant's Petititon for Rehearing *En Banc*' – Motorola *Motorola Mobility LLC v. AU Optronics*, No. 14-8003, 2014 WL 1243797 (7th Cir. Mar. 27, 2014); p. 4. In making this statement, the Brief references Bryant and Eckard (1991) and Combe *et al.* (2008).

Table 1 (and, for simplicity, assume $\beta = \rho$), the fraction of cartels that are not discovered equals $[\lambda(1 - \rho)]/[\lambda(1 - \rho) + \rho]$. As already noted, λ and ρ are not separately identified so it would seem one could not infer how many cartels go undiscovered. If the estimated survival rate $s = (1 - \lambda)(1 - \rho)$ is substituted into this expression, the fraction of cartels that escape discovery can be expressed as $1 - [\rho/(1 - s)]$. Given that $\rho \in (0, 1 - s]$, the fraction of cartels that escape discovery can range from almost none (when ρ is close to $1 - s$) to almost all (when ρ is close to zero). There is then no basis upon which to infer how many cartels go undiscovered.

To summarise the results of this Section, if all cartels are subject to the same birth-death-discovery process then the population of discovered cartels is a representative sample of the population of cartels, which implies:

(*i*) average duration of discovered cartels is an unbiased measure of average cartel duration; and

(*ii*) the estimated probability of death for discovered cartels is an unbiased measure of the probability of death for cartels.

However, the probability of collapse and the probability of discovery cannot be separately identified and this has the implication that:

(*i*) assuming there is no model misspecification, the estimated death rate in studies based on the duration of discovered cartels is an upper bound on the probability a cartel is caught and convicted; and

(*ii*) nothing can be inferred about the fraction of cartels that escape detection.

3. Cartels are Subject to Different Death and Discovery Processes

3.1. *General Properties*

Given that it is clearly restrictive to assume that all cartels are subject to the same stochastic process determining death and detection, let us enrich the environment by allowing cartels to be generated by different processes. This is modelled by allowing the values for (λ, ρ, β) to vary across cartels. This heterogeneity could be due, for example, to industry traits or the characteristics of the cartel. For example, the probability of cartel collapse, λ, could depend on product substitutability and demand variability, both of which affect the gains to cheating and the losses from punishment; or on the ease with which firms can monitor compliance. The discovery parameters ρ and β could depend on the number of firms involved in the cartel as well as the type of consumers they face, where industrial customers are more likely to detect collusion than consumers in a retail market.

In each period, a unit mass of cartels is born and $k : [0, 1]^3 \rightarrow R_+$ is the density function over cartel traits (λ, ρ, β) for those newly formed cartels. By the analysis in the preceding Section, the average cartel duration for a type-(λ, ρ, β) cartel is $CD(\lambda, \rho) \equiv (\lambda + \rho - \lambda\rho)^{-1}$. The average cartel duration across cartels of all types is simply the weighted average of the average cartel duration for a type-(λ, ρ, β) cartel where the weight is $k(\lambda, \rho, \beta)$:

$$\int \int \int (\lambda + \rho - \lambda\rho)^{-1} k(\lambda, \rho, \beta) \, \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta.$$

In calculating the average duration of discovered cartels (or average discovery time), first note that the mass of discovered cartels with duration 1 is:

$$\int \int \int [(1 - \lambda)\rho + \lambda\beta] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta,$$

with duration 2 is:

$$\int \int \int (1 - \lambda)(1 - \rho)[(1 - \lambda)\rho + \lambda\beta] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta,$$

and with duration T is:

$$\int \int \int [(1 - \lambda)(1 - \rho)]^{T-1}[(1 - \lambda)\rho + \lambda\beta] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta.$$

The total mass of discovered cartels is then:

$$\Gamma \equiv \sum_{t=1}^{\infty} \int \int \int [(1 - \lambda)(1 - \rho)]^{t-1}[(1 - \lambda)\rho + \lambda\beta] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta$$

$$= \int \int \int \left[\frac{(1 - \lambda)\rho + \lambda\beta}{\lambda + \rho - \lambda\rho}\right] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta.$$

Hence, average discovery time is:

$$\left(\frac{1}{\Gamma}\right) \sum_{t=1}^{\infty} \int \int \int t[(1 - \lambda)(1 - \rho)]^{t-1}[(1 - \lambda)\rho + \lambda\beta] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta$$

$$= \left(\frac{1}{\Gamma}\right) \int \int \int \left\{\frac{1}{[1 - (1 - \lambda)(1 - \rho)]^2}\right\} [(1 - \lambda)\rho + \lambda\beta] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta$$

$$= \left(\frac{1}{\Gamma}\right) \int \int \int \left(\frac{1}{\lambda + \rho - \lambda\rho}\right) \left[\frac{(1 - \lambda)\rho + \lambda\beta}{\lambda + \rho - \lambda\rho}\right] k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta$$

$$= \left(\frac{1}{\Gamma}\right) \int \int \int CD(\lambda, \rho) P(\lambda, \rho, \beta) CD(\lambda, \rho) k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta,$$

where $P(\lambda, \rho, \beta) \equiv (1 - \lambda)\rho + \lambda\beta$ is the per period probability that a cartel is detected.

The task before us is to compare average duration for all cartels (denoted *ACD* for average cartel duration):

$$ACD \equiv \int \int \int CD(\lambda, \rho) k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta, \tag{1}$$

and average duration for discovered cartels (denoted *ADT* for average discovery time):

$$ADT \equiv \int \int \int \theta(\lambda, \rho, \beta) CD(\lambda, \rho) k(\lambda, \rho, \beta) \mathrm{d}\lambda \mathrm{d}\rho \mathrm{d}\beta, \tag{2}$$

where:

$$
\begin{aligned}
\theta(\lambda, &\rho, \beta) \\
&\equiv \left\{ \int \int \int \left[\frac{(1-\lambda)\rho + \lambda\beta}{\lambda + \rho - \lambda\rho} \right] k(\lambda, \rho, \beta)\mathrm{d}\lambda\mathrm{d}\rho\mathrm{d}\beta \right\}^{-1} \left[\frac{(1-\lambda)\rho + \lambda\beta}{\lambda + \rho - \lambda\rho} \right]. \\
&= \left(\frac{1}{\Gamma} \right) P(\lambda, \rho, \beta)\, CD(\lambda, \rho).
\end{aligned} \tag{3}
$$

In calculating the average cartel duration, the average duration of a type-(λ, ρ, β) cartel is weighted by how many of them are in the population which is $k(\lambda, \rho, \beta)$. However, average discovery time weights the average duration of a type-(λ, ρ, β) cartel by $\theta(\lambda, \rho, \beta)k(\lambda, \rho, \beta)$. If $\theta(\lambda, \rho, \beta)$ is correlated with average duration then average discovery time is a biased measure of average duration.

In examining $\theta(\lambda, \rho, \beta)$, note that it is proportional to the probability that a type-(λ, ρ, β) cartel is discovered in a given period, $P(\lambda, \rho, \beta)$, multiplied by average cartel duration for a cartel of type-(λ, ρ, β), $CD(\lambda, \rho)$. If one thinks of discovery as a sampling process, $\theta(\lambda, \rho, \beta)$ has a natural interpretation. A cartel with a higher probability of discovery is more likely to be sampled and thus will be more heavily weighted in the calculation of average discovery time. A cartel with longer duration has more opportunities of being sampled because there are more periods for which it can be discovered; a cartel that is not discovered in the year of its death avoids being sampled. Thus, higher duration results in a higher value for $\theta(\lambda, \rho, \beta)$ and those cartels are more heavily weighted in the calculation of average discovery time.

Generally, we expect average discovery time to differ from average cartel duration and, in the ensuing analysis, we seek to characterise how they differ. But before doing so, it is worth noting that the two measures are identical when $\beta = 1$.[12] Examining (3), $\beta = 1$ implies $\theta(\lambda, \rho, 1) = 1$ for all (λ, ρ) and, therefore, $ADT = ACD$. The intuition is immediate. If an active cartel is discovered then the act of discovery causes its death so its discovery time equals its duration. If a cartel that collapses is then discovered (as is the case when $\beta = 1$) then again its discovery time equals its duration. In other words, the population of discovered cartels is the same as the population of cartels in which case the 'sample' of cartels derived from discovery is actually the universe of cartels. Clearly, there is no bias in that situation.

When instead $\beta < 1$, average discovery time is generally a biased measure of average cartel duration. The issue is what we can say about the direction and magnitude of that bias: Does average discovery time tend to over or underestimate average cartel duration? Given that ρ and β are apt to be positively related, let us suppose for the moment that $\beta = \varphi(\rho)$ where $\varphi'(\rho) \geq 0$. The weighting factor in ADT is then:

---

[12] A leniency programme could possibly result in β being close to one. As shown in Harrington and Chang (2015), if leniency is full and firms do not anticipate colluding again then, upon collapse, firms will act to minimise expected penalties which implies it is a dominant strategy to apply for leniency. Hence, all dying cartels are discovered.

$$\theta[\lambda, \rho, \varphi(\rho)] = \left(\frac{1}{\Gamma}\right)\left[\frac{(1-\lambda)\rho + \lambda\varphi(\rho)}{\lambda + \rho - \lambda\rho}\right],$$

and is increasing in the probability of discovery:

$$\frac{\partial\theta[\lambda, \rho, \varphi(\rho)]}{\partial\rho} = \left(\frac{1}{\Gamma}\right)\left\{\frac{(1-\lambda)\lambda[1 - \varphi(\rho)] + \lambda\varphi'(\rho)[(1-\lambda)\rho + \lambda]}{(\lambda + \rho - \lambda\rho)^2}\right\} > 0.$$

Not surprisingly, cartels endowed with a higher discovery probability are more likely to be sampled (through discovery) and thus are given more weight in the calculation of *ADT*. Furthermore, holding λ fixed, a higher discovery probability results in shorter cartel duration: $(\lambda + \rho - \lambda\rho)^{-1}$ is decreasing in ρ. Thus, variation in ρ causes cartels with shorter duration to be weighted more, holding λ fixed. Cartels that are more likely to be discovered are more heavily represented in the population of discovered cartels and cartels that are more likely to be discovered have shorter duration. This effect causes average discovery time to underestimate average cartel duration.

Next consider the effect of changing the probability of collapse:

$$\frac{\partial\theta[\lambda, \rho, \varphi(\rho)]}{\partial\lambda} = -\left(\frac{1}{\Gamma}\right)\left\{\frac{\rho[1 - \varphi(\rho)]}{(\lambda + \rho - \lambda\rho)^2}\right\} < 0.$$

Cartels that are more likely to collapse are then underrepresented in the population of discovered cartels because they are more likely to die and disappear before being detected. Thus, variation in λ causes cartels with shorter duration to be weighted less, holding ρ fixed. Cartels that are less likely to collapse are more heavily represented in the population of discovered cartels and cartels that are less likely to collapse have longer duration. This effect causes average discovery time to overestimate average cartel duration.

Summing up to this point, average duration of discovered cartels could either be an overestimate or an underestimate of average duration of all cartels. Relatively unstable cartels (i.e. those with a high value for λ) tend to be under-represented in the pool of discovered cartels because they are more likely to die before being discovered. Given that unstable cartels have shorter duration, this effect will cause average discovery time to overestimate average cartel duration. Cartels that are relatively likely to be discovered (that is, those with a high value for ρ) tend to be over-represented in the pool of discovered cartels. Given that cartels more likely to be discovered have shorter duration then this effect will cause average discovery time to underestimate average cartel duration. The following result is immediate.

PROPERTY 2. *If there is variation in the discovery probability ρ but not in the collapse probability λ then average duration of discovered cartels is an underestimate of average cartel duration: ADT < ACD. If there is variation in the collapse probability λ but not in the discovery probability ρ then average duration of discovered cartels is an over-estimate of average cartel duration: ADT > ACD.*

The bias in using duration on discovered cartels to estimate average cartel duration (or the probability of cartel death) comes from inter-industry variation in the stochastic processes determining death and detection. If there is no variation – all cartels are subject to the same stochastic process – then there is no bias.

3.2. *Numerical Analysis*

The preceding Section showed:

(*i*) when there is variation across cartels in the probability of collapse and/or the probability of discovery then, generically, average discovery time is a biased measure of average cartel duration;

(*ii*) when there is only variation in the probability of collapse then average discovery time is an overestimate of average cartel duration; and

(*iii*) when there is only variation in the probability of discovery then average discovery time is an underestimate of average cartel duration.

It is then variation in cartel collapse and discovery rates that creates bias. A reasonable conjecture is that more variation in collapse (discovery) probabilities ought to make average discovery more of an overestimate (underestimate) of average cartel duration. To assess the validity of that conjecture, we turn to numerical analysis.

To reduce the number of parameters, set $\beta = \rho$ so that the probability of discovery is the same whether the cartel is active or has just collapsed. The task is to specify distributions on the parameters, calculate average cartel duration and average discovery time, and measure bias which we will define as the percentage difference between *ADT* and *ACD*:

$$Bias = 100 \left(\frac{ADT - ACD}{ACD} \right).$$

Note that positive (negative) bias means that average discovery time over-estimates (underestimates) average cartel duration.

Both λ and ρ are assumed to follow beta distributions.[13] A beta distribution is defined by two shape parameters and can be uniquely characterised by the mean μ and coefficient of variation c (which is the standard deviation divided by the mean). We first consider the case where λ and ρ are independently distributed with the following parameters:

$$(\mu_\lambda, \mu_\rho) \in \{0.05, 0.10, 0.15\}^2, \quad (c_\lambda, c_\rho) \in \{0.1, 0.4, 0.8\}^2.$$

Thus, the average probability of collapse or of discovery ranges from 0.05 to 0.15 per period, which correspond to a per period death rate ranging from 0.10 to 0.28. The standard deviation can be small – only 10% of the mean – or large – 80% of the mean. In total, there are 81 parameterisations.

The results are displayed in Table 2. In reading the Tables, consider Table 2 panel (*a*) which reports average cartel duration for each parameterisation. There are nine rows and nine columns. The top three rows correspond to $\mu_\lambda = 0.05$, the middle three rows to $\mu_\lambda = 0.1$, and the bottom three rows to $\mu_\lambda = 0.15$. Analogously, the three columns on the left correspond to $\mu_\rho = 0.05$, the middle three columns to $\mu_\rho = 0.1$, and the three columns on the right to $\mu_\rho = 0.15$. Within each block of three rows and three columns (corresponding to particular means for λ and ρ), there are three rows and columns associated with different values for (c_λ, c_ρ). For example, if

---

[13] Numerical analysis was also conducted using uniform and triangular density functions and the results are qualitatively similar.

Table 2

Average Cartel Durations (a), Average Discovery Time (b) and Biases in Percentage (c) When λ and ρ are Independent (ψ = 1)

Periods	$\mu_\rho = 0.05$			$\mu_\rho = 0.1$			$\mu_\rho = 0.15$		
	$c_\rho = 0.1$	$c_\rho = 0.4$	$c_\rho = 0.8$	$c_\rho = 0.1$	$c_\rho = 0.4$	$c_\rho = 0.8$	$c_\rho = 0.1$	$c_\rho = 0.4$	$c_\rho = 0.8$
Panel (*a*): Average cartel durations									
$\mu_\lambda = 0.05$									
$c_\lambda = 0.1$	10.31	10.66	11.64	6.93	7.39	8.78	5.23	5.68	7.26
$c_\lambda = 0.4$	10.66	11.11	12.42	7.03	7.54	9.2	5.26	5.75	7.57
$c_\lambda = 0.8$	11.64	12.42	15.21	7.3	7.96	10.69	5.37	5.94	8.65
$\mu_\lambda = 0.1$									
$c_\lambda = 0.1$	6.93	7.03	7.3	5.29	5.46	5.95	4.28	4.49	5.13
$c_\lambda = 0.4$	7.39	7.54	7.96	5.46	5.69	6.35	4.36	4.61	5.42
$c_\lambda = 0.8$	8.78	9.2	10.69	5.95	6.35	7.9	4.58	4.95	6.55
$\mu_\lambda = 0.15$									
$c_\lambda = 0.1$	5.23	5.26	5.37	4.28	4.36	4.58	3.62	3.73	4.06
$c_\lambda = 0.4$	5.68	5.75	5.94	4.49	4.61	4.95	3.73	3.88	4.33
$c_\lambda = 0.8$	7.26	7.57	8.65	5.13	5.42	6.55	4.06	4.33	5.51
Panel (*b*): Average discovery time									
$\mu_\lambda = 0.05$									
$c_\lambda = 0.1$	10.3	10.26	9.98	6.92	7.1	7.28	5.22	5.47	5.88
$c_\lambda = 0.4$	11.01	11.09	11.04	7.12	7.38	7.81	5.29	5.6	6.23
$c_\lambda = 0.8$	12.85	13.45	15.11	7.6	8.12	9.75	5.49	5.95	7.54
$\mu_\lambda = 0.1$									
$c_\lambda = 0.1$	6.95	6.83	6.49	5.29	5.26	5.07	4.27	4.31	4.25
$c_\lambda = 0.4$	7.88	7.79	7.48	5.63	5.66	5.6	4.43	4.52	4.6
$c_\lambda = 0.8$	10.71	11.06	11.99	6.56	6.87	7.79	4.85	5.14	6.05
$\mu_\lambda = 0.15$									
$c_\lambda = 0.1$	5.25	5.15	4.9	4.28	4.21	4	3.62	3.59	3.44
$c_\lambda = 0.4$	6.22	6.12	5.82	4.71	4.68	4.52	3.84	3.86	3.79
$c_\lambda = 0.8$	9.71	9.99	10.8	5.97	6.21	6.96	4.46	4.68	5.39
Panel (*c*): Biases in percentage									
$\mu_\lambda = 0.05$									
$c_\lambda = 0.1$	0.0	-3.7	-14.3	-0.1	-3.9	-17.1	-0.2	-3.7	-19.1
$c_\lambda = 0.4$	3.3	-0.2	-11.1	1.2	-2.2	-15.2	0.6	-2.6	-17.7
$c_\lambda = 0.8$	10.4	8.3	-0.6	4.2	2.1	-8.8	2.2	0.1	-12.9
$\mu_\lambda = 0.1$									
$c_\lambda = 0.1$	0.2	-2.9	-11.1	0.0	-3.8	-14.8	-0.1	-4.1	-17.2
$c_\lambda = 0.4$	6.7	3.4	-6.0	3.1	-0.4	-11.8	1.7	-1.9	-15.1
$c_\lambda = 0.8$	22.0	20.2	12.2	10.2	8.2	-1.4	5.8	3.7	-7.6
$\mu_\lambda = 0.15$									
$c_\lambda = 0.1$	0.4	-2.1	-8.8	0.1	-3.3	-12.7	0.0	-3.8	-15.2
$c_\lambda = 0.4$	9.4	6.4	-2.1	5.0	1.6	-8.8	3.0	-0.6	-12.6
$c_\lambda = 0.8$	33.7	32.1	24.9	16.4	14.7	6.2	9.9	8.0	-2.2

$(\mu_\lambda, \mu_\rho) = (0.1, 0.1)$ and $(c_\lambda, c_\rho) = (0.4, 0.8)$ then the cartel duration is 6.35 which means on average a cartel survives for 6.35 periods.

As conjectured, more variation in λ results in a larger over-estimate, while more variation in ρ results in a larger underestimate. To show the former result, hold fixed the distribution on ρ and the mean of λ. As the inter-industry variation in λ is increased (which occurs when the coefficient of variation increases holding the mean fixed), the bias increases. For example, if $(\mu_\lambda, \mu_\rho) = (0.1, 0.1)$ and $c_\rho = 0.1$ (so there is little variation in ρ) then increasing variation in λ by raising c_λ from 0.1 to 0.4 to 0.8 causes the bias to go from 0.0% (bias is almost zero) to 3.1% (average discovery time slightly over-estimates average cartel duration) to 10.2% (average discovery time more significantly over-estimates average cartel duration). In all cases, as c_λ increases – holding μ_λ, μ_ρ, and c_ρ fixed – average discovery time increases relative to average cartel duration so the former is more of an overestimate. Analogously, as c_ρ increases – holding μ_λ, μ_ρ, and c_λ fixed – average discovery time declines relative to average cartel duration so the former is more of an underestimate.

PROPERTY 3. *As the variation across cartels in the probability of collapse increases, the difference between average discovery time and average cartel duration increases so average discovery time is more of an overestimate. As the variation across cartels in the probability of discovery increases, the difference between average discovery time and average cartel duration decreases so average discovery time is more of an underestimate.*

A second takeaway is that the extent of over or underestimation is not very large in magnitude. For most parameterisations, it is less than 20%. It is possible, however, for the absolute size of bias to be significant. This occurs when the mean and variance of the collapse rate is large and the mean and variance of the discovery rate is small. For example, if $(\mu_\lambda, \mu_\rho) = (0.15, 0.05)$ and $(c_\lambda, c_\rho) = (0.8, 0.1)$ then average discovery time over-estimates average cartel duration by about 34%.

Although the case when λ and ρ are independent is a useful benchmark, it could be more natural for λ and ρ to be positively correlated. For example, in an industry with more firms, the probability of collapse may tend to be higher because there is a larger incremental gain to cheating which makes the cartel less stable. In addition, more firms may raise the probability of discovery because there are more people involved so there is a greater chance that someone, either intentionally or inadvertently, would reveal information about the cartel to a customer or the authorities. Furthermore, a higher discovery probability will reduce the expected payoff from collusion and that could lead to a higher probability of collapse.

To assess the robustness of our results, let us then allow λ and ρ to be positively correlated. Using a system of bivariate distributions introduced by Plackett (1965), joint distributions between λ and ρ are constructed that have the same marginal distributions assumed for the independent case. Plackett's system is a general method to generate a bivariate distribution given any pair of marginal distributions. The joint distribution is characterised by one parameter $\psi \geq 0$. As ψ goes from 0 to 1, the two random variables go from being highly negatively correlated to independent. As ψ goes from 1 to $+\infty$, the two random variables are increasingly positively correlated. Hence, one can numerically search over ψ for desired correlation. For illustration, Figure 1 plots the marginal
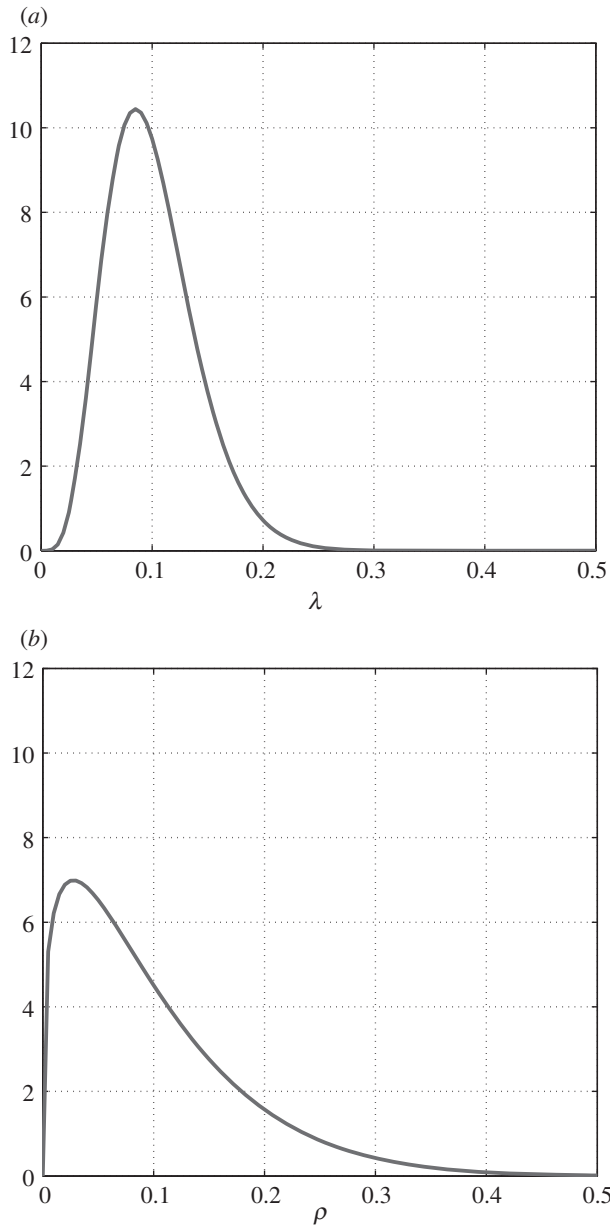
(a)

(b)

Fig. 1. *Marginal Densities.* $\mu_\lambda = \mu_\rho = 0.1$, $c_\lambda = 0.4$, $c_\rho = 0.8$

densities when $(\mu_\lambda, \mu_\rho) = (0.1, 0.1)$ and $(c_\lambda, c_\rho) = (0.4, 0.8)$, and Figure 2 plots the corresponding joint densities for correlation being 0, 0.3 and 0.7 respectively.

Results are in Table 3 for when the correlation between λ and ρ is 0.70 ($\psi \simeq 16$); they are presented in the same manner as Table 2.[14] We find the relationship between

---

[14] The analysis has also been done for lower values of the correlation and the qualitative findings are the same.
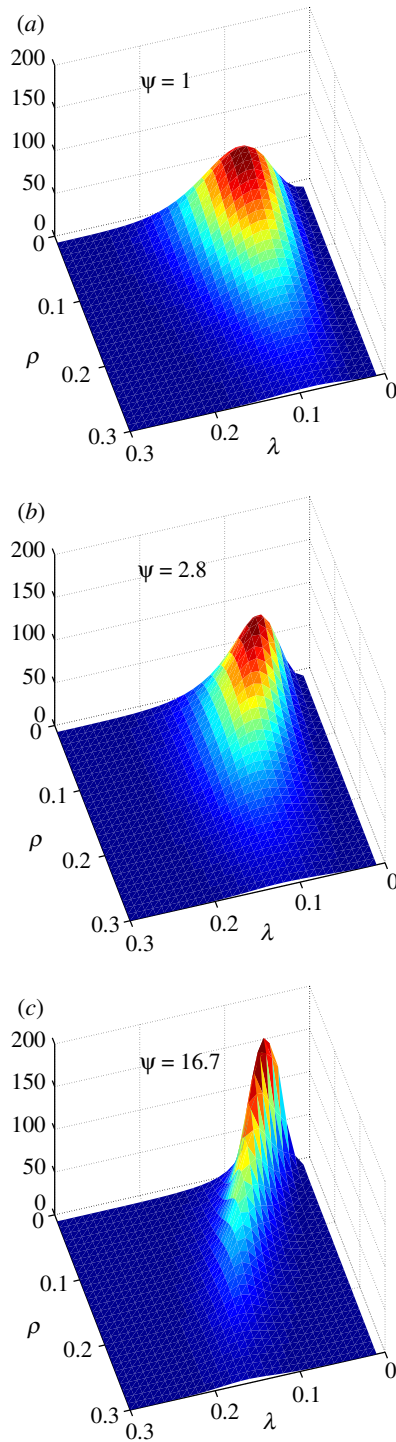
(a)

(b)

(c)

Fig. 2. *Joint Densities, Correlation Being 0, 0.3 and 0.7 Respectively*

Table 3

Average Cartel Durations (a), Average Discovery Time (b) and Biases in Percentage (c) When Correlation is 0.7 ($\psi \simeq 16$)

Periods	$\mu_\rho = 0.05$			$\mu_\rho = 0.1$			$\mu_\rho = 0.15$		
	$c_\rho = 0.1$	$c_\rho = 0.4$	$c_\rho = 0.8$	$c_\rho = 0.1$	$c_\rho = 0.4$	$c_\rho = 0.8$	$c_\rho = 0.1$	$c_\rho = 0.4$	$c_\rho = 0.8$
Panel (a): Average cartel durations									
$\mu_\lambda = 0.05$									
$c_\lambda = 0.1$	10.34	10.82	11.99	6.95	7.49	9.08	5.24	5.75	7.53
$c_\lambda = 0.4$	10.82	11.84	14.2	7.12	7.99	10.69	5.32	6.05	8.87
$c_\lambda = 0.8$	11.99	14.2	22.09	7.47	8.89	15.67	5.47	6.53	12.87
$\mu_\lambda = 0.1$									
$c_\lambda = 0.1$	6.95	7.12	7.47	5.31	5.54	6.13	4.29	4.56	5.3
$c_\lambda = 0.4$	7.49	7.99	8.89	5.54	6.06	7.29	4.42	4.91	6.31
$c_\lambda = 0.8$	9.08	10.69	15.67	6.13	7.29	11.92	4.7	5.62	10.12
$\mu_\lambda = 0.15$									
$c_\lambda = 0.1$	5.24	5.32	5.47	4.29	4.42	4.7	3.63	3.79	4.18
$c_\lambda = 0.4$	5.75	6.05	6.53	4.56	4.91	5.62	3.79	4.14	5
$c_\lambda = 0.8$	7.53	8.87	12.87	5.3	6.31	10.12	4.18	5	8.76
Panel (b): Average discovery time									
$\mu_\lambda = 0.05$									
$c_\lambda = 0.1$	10.34	10.39	10.12	6.95	7.21	7.45	5.24	5.55	6.04
$c_\lambda = 0.4$	11.19	11.8	11.93	7.23	7.89	8.78	5.36	5.96	7.11
$c_\lambda = 0.8$	13.33	15.81	21.8	7.82	9.32	14.36	5.62	6.7	11.26
$\mu_\lambda = 0.1$									
$c_\lambda = 0.1$	6.96	6.85	6.48	5.3	5.32	5.14	4.29	4.37	4.33
$c_\lambda = 0.4$	7.97	8.07	7.58	5.72	6.02	6.02	4.5	4.84	5.07
$c_\lambda = 0.8$	11.16	13.22	17.28	6.8	8.1	11.59	5.01	5.98	9.17
$\mu_\lambda = 0.15$									
$c_\lambda = 0.1$	5.24	5.14	4.84	4.29	4.24	4.02	3.63	3.63	3.48
$c_\lambda = 0.4$	6.25	6.19	5.64	4.78	4.91	4.69	3.9	4.1	4.06
$c_\lambda = 0.8$	10.15	12.12	16.09	6.21	7.45	10.61	4.62	5.54	8.36
Panel (c): Biases in percentage									
$\mu_\lambda = 0.05$									
$c_\lambda = 0.1$	0.0	−4.0	−15.6	−0.1	−3.8	−17.9	−0.1	−3.5	−19.7
$c_\lambda = 0.4$	3.4	−0.3	−16.0	1.5	−1.2	−17.9	0.9	−1.5	−19.8
$c_\lambda = 0.8$	11.1	11.4	−1.3	4.8	4.9	−8.4	2.7	2.6	−12.6
$\mu_\lambda = 0.1$									
$c_\lambda = 0.1$	0.0	−3.7	−13.2	0.0	−4.1	−16.2	−0.1	−4.1	−18.3
$c_\lambda = 0.4$	6.4	1.1	−14.7	3.3	−0.7	−17.4	2.0	−1.4	−19.6
$c_\lambda = 0.8$	23.0	23.6	10.3	10.9	11.1	−2.8	6.4	6.4	−9.4
$\mu_\lambda = 0.15$									
$c_\lambda = 0.1$	0.1	−3.4	−11.5	0.0	−3.9	−14.5	−0.1	−4.2	−16.8
$c_\lambda = 0.4$	8.7	2.3	−13.5	4.9	0.1	−16.5	3.1	−1.1	−18.9
$c_\lambda = 0.8$	34.9	36.7	25.0	17.2	18.0	4.9	10.6	10.8	−4.5

the inter-industry variation of λ and ρ and bias is quantitatively similar to that in the independent case. However, when either λ or ρ have a large degree of variation, there can be a violation of the pattern described in Property 3. For example, when $(\mu_\lambda, \mu_\rho) = (0.1, 0.1)$ and $c_\rho = 0.8$ in Table 3, the bias decreases as the variation in λ is increased by raising c_λ from 0.1 to 0.4. However, these violations are small in magnitude and we believe that the main message of Property 3 is broadly relevant even when λ and ρ are positively correlated.

4. Estimating Cartel Duration and the Probability of Death

Thus far, we have provided a framework within which to think about bias in estimates of cartel duration due to a potentially biased sample in the form of discovered cartels. It was shown that variation in the probability of collapse across cartels results in average duration of discovered cartels being an over-estimate, while variation in the probability of discovery across cartels results in average duration of discovered cartels being an over-estimate. We now turn to estimating the theoretical model and do so with two objectives. First, to determine whether the data are consistent with the presence of bias. We find that there is bias and, though the estimates prove not to shed light on its direction, it does offer bounds on the extent of bias. Second, to provide an estimate of the probability of cartel death that takes account of that bias. Previous approaches deliver an estimate of the probability of cartel death conditional on the cartel being discovered. The theoretical framework of this article is used to estimate the unconditional probability of death, which is the object of interest. However, there are some qualifications to be made with this approach, which we discuss below.

To make the empirical analysis more manageable, it is assumed $\beta = \rho$.[15] Given a density function $k(\cdot)$ for (λ, ρ), the theoretical framework provides a density function on the duration of discovered cartels and on the duration of all cartels. The probability density on duration τ conditional on the cartel being discovered is:

$$\int \int \left\{ \frac{\rho[(1-\lambda)(1-\rho)]^{\tau-1}}{\int \int \left(\frac{\rho}{\lambda + \rho - \lambda\rho} \right) k(\lambda, \rho) \mathrm{d}\lambda \mathrm{d}\rho} \right\} k(\lambda, \rho) \mathrm{d}\lambda \mathrm{d}\rho, \tag{4}$$

and the probability density on duration τ for all cartels is:

$$\int \int (\lambda + \rho - \lambda\rho)[(1-\lambda)(1-\rho)]^{\tau-1} k(\lambda, \rho) \mathrm{d}\lambda \mathrm{d}\rho. \tag{5}$$

The empirical strategy is to specify a functional form on $k(\cdot)$, fit (4) to the data on discovered cartels to recover information on the parameters of $k(\cdot)$, and to use those estimated parameters to simulate the density function on duration for all cartels in (5).

One weakness of this approach is that it is reliant on the functional form specified for $k(\cdot)$. A second weakness is that not all of the parameters of $k(\cdot)$ are identified.

---

[15] As stated in Section 1, leniency programmes could result in $\beta > \rho$. Our empirical analysis will focus on the pre-leniency programme era in the US in which case $\beta = \rho$ is a plausible assumption.

Because of these concerns, we have taken a Bayesian approach which avoids focusing on any particular values of estimates and instead has the more modest goal of providing more informed beliefs over the variables of interest. In light of these qualifications, the ensuing estimates should be taken as an improvement over previous estimates rather than definitive estimates of cartel duration and the probability of cartel death.

In specifying a functional form for $k(\cdot)$, we use the specification in Section 3; hence, a maintained hypothesis is that a cartel draws (λ, ρ) from a joint beta distribution defined by the five parameters $\theta \equiv (\mu_\lambda, \mu_\rho, c_\lambda, c_\rho, \psi)$. A Bayesian approach requires specifying a prior set of beliefs on θ which, reflecting our limited prior information, is assumed to be a uniform distribution with a fairly wide support. The next step is to use the data on discovered cartels to inform us about θ. More specifically, given a particular value for θ, we calculate the likelihood of observing the duration data on discovered cartels for when the density function takes the form in (4). Using this likelihood, the prior distribution on θ is updated to derive a posterior distribution on θ. Based on these posterior beliefs on θ, a distribution can be derived on the cartel death rate and on the bias associated with using average discovered cartel duration as a proxy for average cartel duration.

4.1. *Data*

Before turning to a more detailed description of the empirical method, it is useful to describe the data. Coming from the Commerce Clearing House Trade Regulation Reports, our sample consists of price-fixing cases pursued by the Antitrust Division of the US Department of Justice which resulted in a conviction, guilty plea, or *nolo contendere*. This data source is the same as that used in Bryant and Eckard (1991) whose data set encompassed 184 indicted cartels that were discovered over the period 1961–88.[16] Their data set has been extended so that it now runs from 1961 to 2004 and includes 429 discovered cartels.[17] As explained below, a proper implementation of our empirical strategy requires using only a subset of the data.

4.2. *Empirical Method*

In each year, a collection of cartels are born and each member of that cohort of cartels will eventually have a realised duration. In the subsequent discussion, τ refers to a cartel's duration and T to a particular year. Suppose that we observe cartels which were discovered (and, therefore, died) between years T_1 and T_2. For each of those cartels, there is a birth year and a duration. Organise the data by birth cohort and let $N_{\tau,T}$

---

[16] We follow the same data criteria as in Bryant and Eckard (1991). All cartels listed in the Topical Index for Antitrust Cases under 'price fixing' are included. Duration is defined in the same way as their DUR2. Specifically, duration is from the cartel's latest start date to its earliest discovery date where duration is rounded to the nearest greatest number of years. The indictment date is used if the discovery date is not available. We chose to use DUR2 rather than their alternative measure DUR1 because the latter has a mode of cartel duration well above one year which is less consistent with our model.
[17] In 1995, CCH was bought by Wolters Kluwer Law & Business which now provides the CCH data online. Unfortunately, the Topical Index for Antitrust Cases is outdated and does not include cases after 2004.

denote the number of cartels born in year T that had duration τ in the data. While there is no bound on the earliest cohort one could have represented in the data, the latest possible cohort is T_2 which is associated with discovering a cartel with duration 1 in year T_2. For a cohort $T < T_1$, only cartels of duration $T_1 - T + 1$ to $T_2 - T + 1$ can be observed; cartels of duration less than $T_1 - T + 1$ would have been discovered prior to the initial year in which discoveries were recorded. For cohort $T \in \{T_1, T_1 + 1, \ldots, T_2\}$, only cartels of duration 1 to $T_2 - T + 1$ can be observed. The data for cohort T is then represented by the vector $N_{*,T} \equiv (N_{\max\{1, T_1 - T + 1\}, T}, \ldots, N_{T_2 - T + 1, T})$, and we will have $N_T \equiv \sum_{\tau = \max\{1, T_1 - T + 1\}}^{T_2 - T + 1} N_{\tau, T}$ denote the number of discovered cartels from cohort T. The data set is then the collection of vectors $N_{*,T}$ for all cohorts with at least one observation. In the ensuing description, we will let Ω denote the data set and Λ denote the set of cohorts.

We want to compute the likelihood of observing Ω if each cartel's value for (λ, ρ) is drawn from a joint beta distribution given θ. To compute the likelihood, consider cohort T (that is, discovered cartels that were born in year T). The probability of observing the cohort T vector of durations $N_{*,T}$ conditional on there being N_T discovered cartels is, from the theoretical model:

$$\Pr(N_{*,T}|N_T, \theta) = \binom{N_T}{N_{\max\{1, T_1 - T + 1\}, T}, \ldots, N_{T_2 - T + 1, T}} \prod_{\tau = \max\{1, T_1 - T + 1\}}^{T_2 - T + 1} \left[\frac{p(\tau, \theta)}{Q_T(\theta)}\right]^{N_{\tau, T}}, \quad (6)$$

where $\binom{N_T}{N_{\max\{1, T_1 - T + 1\}, T}, \ldots, N_{T_2 - T + 1, T}}$ is the multinomial coefficient, and

$$Q_T(\theta) = \sum_{\tau = \max\{1, T_1 - T + 1\}}^{T_2 - T + 1} \int \int \rho[(1 - \lambda)(1 - \rho)]^{\tau - 1} k(\lambda, \rho) \mathrm{d}\lambda \mathrm{d}\rho, \quad (7)$$

is the probability of being discovered with duration between $\max\{1, T_1 - T + 1\}$ and $T_2 - T + 1$, and:

$$p(\tau, \theta) = \int \int \rho[(1 - \lambda)(1 - \rho)]^{\tau - 1} k(\lambda, \rho) \, \mathrm{d}\lambda \mathrm{d}\rho, \quad (8)$$

is the probability of being discovered with duration τ. In other words, if there are N_T independent draws for which the likelihood of a draw taking value τ is $p(\tau, \theta)/Q_T(\theta)$ for $\tau \in \{\max\{1, T_1 - T + 1\}, \ldots, T_2 - T + 1\}$ (and 0 otherwise) then the probability of observing $N_{*,T}$ is (6). Given (6) is the likelihood of observing the vector of durations for cohort T, conditional on their being N_T discovered cartels, the probability of observing data from cohorts in Λ, conditional on the number of discovered cartels in each year, is:

$$\Pr(\Omega|\theta, \{N_T\}_{T \in \Lambda}) = \prod_{T \in \Lambda} \Pr(N_{*,T}|N_T, \theta). \quad (9)$$

To use these calculated likelihoods in a Bayesian framework, a prior distribution on the population distribution parameters θ is specified which is assumed to be the uniform conditional on the number of discoveries:

$$\Pr(\theta|\{N_T\}_{T \in \Lambda}) \sim U(\Theta). \quad (10)$$

Our estimation does not then use any information from the number of discovered cartels and is based solely on the duration of those discovered cartels.[18] The difficulty in deriving any information from the number of discovered cartels is that it requires some understanding about how many cartels there are which means specifying a cartel birth process. At this stage, we have little insight into the cartel birth process. Furthermore, if the number of discovered cartels were to be used then a prior distribution on the number of cartels would have to be specified and, even if a uniform prior is used, results would be highly sensitive to the specified support. For these reasons, the number of discovered cartels was not used in our estimation.

We next specify the support of the uniform prior distribution on $\theta \equiv (\mu_\lambda, \mu_\rho, c_\lambda, c_\rho, \psi)$. The means of the distributions have support: $(\mu_\lambda, \mu_\rho) \in [0.025, 0.20]^2$, which implies the annual probability of death can range from 0.049 to 0.36. The coefficients of variation are assumed to have supports:

$$c_\lambda \in \left[0, \sqrt{\frac{1-\mu_\lambda}{1+\mu_\lambda}}\right] \quad \text{and} \quad c_\rho \in \left[0, \sqrt{\frac{1-\mu_\rho}{1+\mu_\rho}}\right].$$

The lower bound of zero allows for the homogeneous case (so all cartels have the same λ and ρ), while the upper bound is chosen to preclude the case when the density $k(\lambda, \rho)$ is positive at $(\lambda, \rho) = (0, 0)$ which would mean average cartel duration is infinity.[19] Finally, $\psi \in [1,10]$, which allows the correlation to range from 0 to approximately 0.6. We also experimented with other reasonable parameter space specifications and did not find any significant change in the results.[20] Table 4 reports the first two moments of the prior distribution. Based on this distribution, the mean annual cartel death rate of 0.21 and the range based on two standard deviations is

Table 4

Moments of Prior and Posterior Distributions on θ

	Prior mean	Prior SD	Posterior mean	Posterior SD
μ_λ	0.110	0.051	0.088	0.037
μ_ρ	0.110	0.051	0.092	0.038
c_λ	0.449	0.261	0.299	0.240
c_ρ	0.450	0.260	0.253	0.222
ψ (correlation)	5.5 (0.49)	2.6 (0.13)	5.3 (0.48)	2.6 (0.13)
Death rate	0.207	0.062	0.174	0.015

Notes. There are 224 data observations. The moments are estimated from over 2,500 draws from the distribution for birth cohorts 1960–85.

[18] Similarly, Bryant and Eckard (1991) used only duration data to estimate the death rate. Though they did use the number of discoveries to estimate the birth rate, that estimation required making the assumption that all cartels are discovered and it is the problematic nature of such an assumption that is the starting point to this article.

[19] When $c_\lambda > \sqrt{(1-\mu_\lambda)/(1+\mu_\lambda)}$, the alpha of the marginal for the beta distribution on λ is larger than 1, which implies that the density is positive at 0. The same applies to c_ρ.

[20] We considered $(\mu_\lambda, \mu_\rho) \in [0.05, 0.175]^2$ and $(\mu_\lambda, \mu_\rho) \in [0.05, 0.15]^2$, $(c_\lambda, c_\rho) \in [0, 0.75\sqrt{(1-\mu_\lambda)/(1+\mu_\lambda)}] \times [0, 0.75\sqrt{(1-\mu_\rho)/(1+\mu_\rho)}]$, and $\psi = 1$.

[0.084,0.336]. In specifying the prior distribution, we want to have highly dispersed beliefs so that the data on discovered cartel duration – not the prior on θ – largely drives the posterior beliefs on θ.

Given (9) and (10), the posterior distribution on the population parameter vector θ is:

$$\Pr(\theta|\Omega) \propto \Pr(\Omega|\theta, \{N_T\}_{T\in\Lambda})\Pr(\theta|\{N_T\}_{T\in\Lambda})$$

$$\propto \prod_{T\in\Lambda} \prod_{\tau=\max\{1,T_1-T+1\}}^{T_2-T+1} \left[\frac{p(\tau,\theta)}{Q_T(\theta)}\right]^{N_{\tau,T}} \tag{11}$$

$$= \prod_{\tau=\max\{1,T_1-T+1\}}^{T_2-T+1} p(\tau,\theta)^{N_{\tau,T}} \Big/ \prod_{T\in\Lambda} Q_T(t,\theta)^{N_T}.$$

It is with this posterior distribution that we can provide measures of the cartel death rate and the possible extent of bias from using discovered cartel duration as a proxy for cartel duration. For each value of $\theta \in \Theta$, the mean cartel death rate can be calculated in a straightforward manner. The posterior beliefs on θ in (11) will then give us posterior beliefs on the mean cartel death rate. Analogously, for each value of $\theta \in \Theta$, bias can be calculated using the method in Section 3 and the posterior beliefs on θ are then used to generate posterior beliefs on bias.

This empirical strategy requires that the distribution on (λ, ρ) is fairly stable over time.[21] However, there was a potentially radical change in θ with the revision of the leniency programme in 1993.[22] For this reason, we have restricted our analysis to cartels that are largely from the pre-leniency programme era.[23] The ensuing analysis is based on cartel cohorts starting in 1960 on through 1985, so we only use the subset $\{1960, \ldots, 1985\}$ of Λ. 1960 is a natural starting date as it is the year the electrical equipment conspiracy was indicted which could have signalled a new regime in terms of enforcement. Given we have data on cartels discovered during 1961–2004, durations in $\{1, \ldots, 2005 - T\}$ are observed for cohort T with the exception of $T = 1960$ for which we observe durations in $\{2, \ldots, 45\}$. The 26 cohort years covered by 1960–85 encompass 224 cartels. While six of those cartels were discovered after 1993, this means less than 3% of the observations are from the leniency programme era. Figure 3 provides the histogram on duration for the discovered cartels used in the empirical analysis. The mean duration for discovered cartels is 5.8 years with a fairly substantial standard deviation of 4.7 years.[24]

While a separate analysis could be done using cartel cohorts during (and just before) the leniency programme era, many of the cohorts will be censored because data are

---

[21] The analogous maintained assumption in previous empirical analyses is that (λ, ρ) is stable over time.

[22] For information on leniency programmes, see Spagnolo (2008); and for details on the programme operated by the US Department of Justice's Antitrust Division, see Hammond (2005).

[23] A change in enforcement policy will result in a change in the stochastic process on birth, death and discovery which results in a transition to a new stationary distribution on cartels. These dynamics are explored in Harrington and Chang (2009) and some results are offered for drawing inferences about the impact of a change in policy. Miller (2009) pursues this type of approach to assess empirically the impact of the 1993 revision.

[24] These moments are very close to those for the entire data set of 429 cartels discovered over 1961–2004, which has a mean of 5.4 years and a standard deviation of 4.7 years.
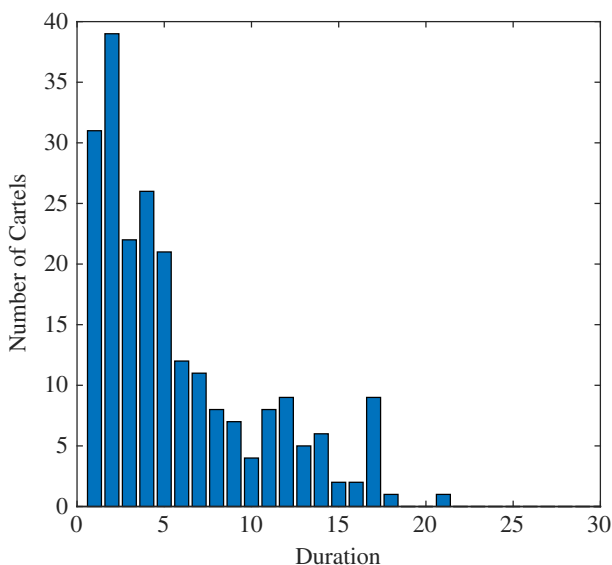
Fig. 3. *Empirical Distribution on Duration for Discovered Cartels, Cohorts 1960–85*
Note. Mean = 5.8 years, standard deviation = 4.7 years.

available only for cartels discovered through 2004. For example, we only observe durations of one, two and three years for the cohort of cartels born in 2002. Though the empirical model takes such censoring into account when it calculates the likelihood, the problem is a lack of information due to censoring. We conducted the analysis for cohorts running from 1985 to 2004, of which 64% of the cartels were discovered post-1993. The likelihood function based on (4) is essentially flat; that is, the likelihood of observing the durations for these cohorts of discovered cartels is very similar for almost all values of θ in the support. Given there is very little information in the data to discriminate among values of θ, the posterior beliefs on θ are basically the same as the prior beliefs. Further analysis was able to attribute this finding to the censoring of cartels with long duration. Estimating the model for cohorts 1960–85 after excluding cartels with more than 10 years of duration, the posterior beliefs were quite close to the prior beliefs. Furthermore, this property is not due to a reduction in the number of observations. If we instead estimated the model using only data from even-year cohorts over 1960–85 – thereby reducing the number of observations from 224 to 121 (which is a far larger reduction than the 43 observations lost by excluding cartels whose duration exceeds 10 years) – the estimates are very similar to when we use all cohorts over 1960–85.

4.3. *Estimates*

The moments for the posterior distribution on $\theta \equiv (\mu_\lambda, \mu_\rho, c_\lambda, c_\rho, \psi)$ are reported in Table 4. The first point to note is that the posterior beliefs are clearly differ from the prior beliefs. The prior mean of λ and ρ are 0.110, while the posterior means are 0.088 and 0.092 respectively. The means of the coefficient of variations are significantly

reduced from 0.450 to 0.299 for λ and to 0.253 for ρ. However, there is effectively no change in beliefs on ψ. What all this means is that, within the specified support on θ, the data on discovered cartel duration are more consistent with joint distributions on (λ, ρ) which give more mass to low values of (λ, ρ) and to less variation in (λ, ρ) across cartels. The data shed no light on the correlation between λ and ρ (which, recall, is determined by ψ).

To provide more information on the posterior beliefs, Figure 4 plots the joint density functions for a pair of parameters from $(\mu_\lambda, \mu_\rho, c_\lambda, c_\rho, \psi)$ while setting the other three parameter values equal to their posterior means. Figure 4a reports the joint density function on the mean of the probability of discovery and the mean of the probability of collapse, (μ_ρ, μ_λ), when $(c_\lambda, c_\rho, \psi) = (0.299, 0.253, 5.3)$. The 'mountain range' encompasses pairs (μ_ρ, μ_λ) for which the mean death rate is around 0.17 and reflects that we cannot separately identify μ_ρ and μ_λ.[25] Figure 4b reports the joint density function on the coefficients of variation of the probability of collapse and the probability of discovery, (c_λ, c_ρ), when $(\mu_\lambda, \mu_\rho, \psi) = (0.088, 0.092, 5.3)$. The density peaks at $(c_\lambda, c_\rho) = (0, 0)$ which is when there is no variation in (λ, ρ). While there is no bias in that case, note that posterior beliefs assign significant probability to $(c_\lambda, c_\rho) > (0, 0)$ in which case there is bias. Results on bias are provided below. We also find that there is more variation across cartels in the probability of collapse than in the probability of discovery: the mean of c_λ exceeds the mean of c_ρ. Furthermore, this ordering is robust to the different assumptions on the support of θ mentioned earlier. Finally, Figure 4c sets $(\mu_\lambda, \mu_\rho) = (0.088, 0.092)$ and reports the density function on (c, ψ) where $c(= c_\lambda = c_\rho)$ is a common coefficient of variation. There is not much to report here other than that beliefs on ψ are fairly uniform.

Given this posterior distribution on λ and ρ, we can use it to measure the amount of bias associated with measuring cartel duration with data on discovered cartels. For a given distribution on (λ, ρ) – that is, particular values for $(\mu_\lambda, \mu_\rho, c_\lambda, c_\rho, \psi)$ – we showed in Section 3 how one can calculate average duration of discovered cartels, *ADT*, and average duration of cartels, *ACD*. Recall that duration bias is measured by:

$$Bias = 100\left(\frac{ADT - ACD}{ACD}\right).$$

Given posterior beliefs on $(\mu_\lambda, \mu_\rho, c_\lambda, c_\rho, \psi)$, the numerical method for estimating bias is to draw values for $(\mu_\lambda, \mu_\rho, c_\lambda, c_\rho, \psi)$ repeatedly according to those posterior beliefs and, for each draw, calculate *Bias*. In this manner, a density function on bias is constructed.

The solid line in Figure 5 is that posterior density function on *Bias*. The density function peaks at *Bias* = 0 because the posterior density on θ is highest at $(c_\lambda, c_\rho) = (0, 0)$ which, from the analysis in Section 2, implies no bias. In addition, *Bias* can be close to zero even when $(c_\lambda, c_\rho) > (0, 0)$. Recall from Section 3 that variation in λ causes positive bias and variation in ρ causes negative bias. Thus, if there are comparable amounts of variation in both the collapse and discovery probabilities, they will tend to cancel each other out so that bias is small. It is then not surprising that

[25] This is one of the reasons why we adopt a Bayesian approach. The posterior distribution simply reflects our updated beliefs on the parameters given the data and model. Had we adopted a maximum likelihood approach, there would not have been a single point estimate of the parameters.
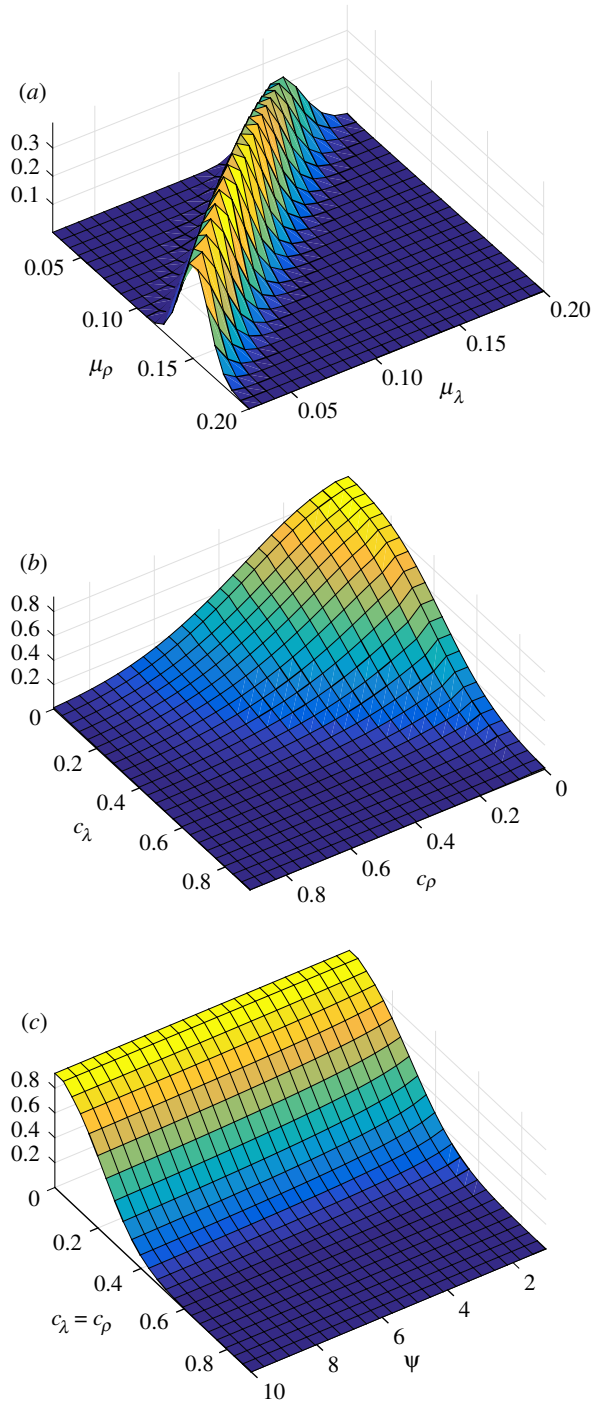
Fig. 4. *Posterior Density Function on* θ
Notes. The graphs show the posterior density as proportion of the density at the mode. For each graph, the unplotted parameters are set equal to their posterior means.
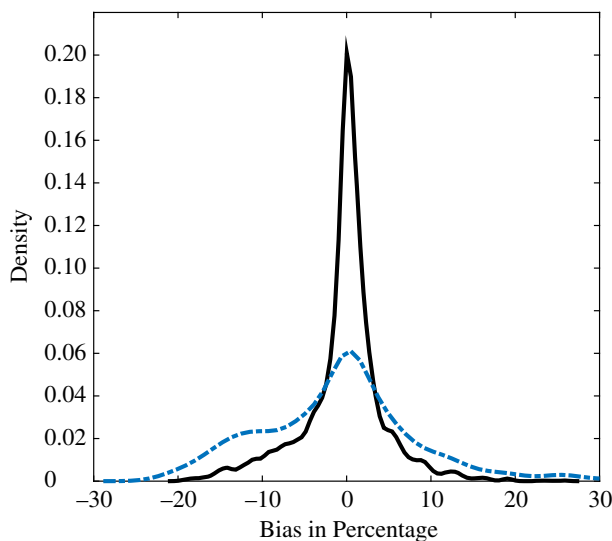
Fig. 5. *Prior and Posterior Density Functions on Bias (in Percentage)*
Notes. These are kernel estimates from 2,500 draws with optimal bandwidth. Solid (dashed) curve is the posterior (prior) density function. Mean of the posterior (prior) density function is −0.30% (−0.29%). Standard deviation of the posterior (prior) density function is 4.7% (10.9%).

the density function on *Bias* peaks at zero. In examining the bounds on bias, posterior beliefs assign positive density ranging from around −15% to +10%. That is, average duration of discovered cartels could overestimate average cartel duration by as much as 10% or underestimate it by as much as 15%. Given average discovered cartel duration in the data set of 5.8 years, this implies a range from 5.27 to 6.82 years.[26]

For comparison, the prior density function on *Bias* is also plotted which is the dashed line. These prior beliefs come from the model specification and a uniform distribution on the parameters. Any difference between the prior and posterior beliefs is driven by the data. As is easily observed, the posterior distribution is much less dispersed than the prior distribution. The standard deviation on the prior density function is 10.9%, while it is only 4.7% on the posterior density function. Thus, if we consider two standard deviations from the mean *Bias* of −0.3%, the range of *Bias* for the prior density function is −22.1% to +21.5% and for the posterior density function is −9.7% to +9.1%. That the posterior distribution is much tighter than the prior distribution is robust to changing the support on the prior distribution. In sum, we find there to be a modest amount of bias associated with using the average duration of discovered cartels as a proxy for the average duration of all cartels.

Finally, and perhaps of primary interest, the posterior distribution on λ and ρ can be used to yield an estimate of the cartel death rate. We construct the average annual probability of death as follows. Given particular values for λ and ρ, a cartel dies with probability $[1 - (1 - \lambda)(1 - \rho)]$ in any period. For a given value for θ, the population average probability of death is then:

---

[26] It is important to keep in mind that there is also sampling error associated with using discovered cartels.

$$\int \int [1 - (1 - \lambda)(1 - \rho)] k(\lambda, \rho; \theta) \mathrm{d}\lambda \mathrm{d}\rho. \tag{12}$$

We then use the posterior beliefs on θ to derive an expectation of the population average probability of death by integrating (12) with respect to those posterior beliefs.[27] This method is designed to take account of the bias associated with using data on discovered cartels. As reported in Table 4, the average annual probability that a cartels dies is 0.174. With a standard deviation of 0.015, a two standard deviation interval around the mean ranges from 0.144 to 0.204, which contains previous estimates. Given our estimates of bias are modest, such a finding is not surprising. It would then seem that, in spite of concerns of bias from using the duration of discovered cartels, we do not find evidence that it has produced particularly misleading estimates.[28]

5. Concluding Remarks

Accurate measures of cartel duration and the probability of cartel death are critical components in assessing to what extent cartels are stable and in evaluating the effect of competition policy in terms of destabilising cartels. At present, there is much confusion and misinformation about what one can infer about the stability of cartels using data on discovered cartels. One of the contributions of this article is constructing a theoretical framework within which to explore sources of bias associated with using discovered cartel duration as a proxy for cartel duration and, hopefully, bring clarity to these issues.

Our theoretical analysis showed that bias is tied to variation in the probabilities of death and discovery across cartels. If that variation is small then bias is small. If variation in the probability of collapse across cartels is large relative to variation in the probability of discovery then average duration of discovered cartels is an over-estimate of average cartel duration; while if variation in the probability of collapse across cartels is small relative to variation in the probability of discovery then average duration of discovered cartels is an underestimate of average cartel duration. This theoretical construct is then used to estimate the extent of bias and deliver an estimate of the probability of death that takes account of that bias. Data on discovered cartel duration is used to recover information on the parameters defining the distribution on the probabilities of collapse and discovery across cartels. With a posterior set of beliefs on those parameters, the distribution on duration for all cartels can be simulated. There is evidence of modest bias associated with using average discovered cartel duration as a proxy for average cartel duration; on the order of 10–15%. Finally, we find that a cartel born between 1960 and 1985 had a 17% chance of dying in a year, either through internal collapse or discovery by the authorities.

This article shows that progress can be made in learning about the latent universe of cartels using data on the observed sample of discovered cartels. The approach is rooted

---

[27] These integrals are approximated by taking 2,500 draws according to the relevant distribution.

[28] For legal manufacturing cartels in Finland during 1951–90, Hyytinen *et al.* (2014*a*) use a hidden Markov birth-death model to estimate an annual probability of cartel birth of 0.22 and an annual probability of cartel death of 0.04. Their estimate of cartel death is far smaller than our estimates for unlawful cartels.

in specifying a class of stochastic processes describing cartel birth, death, and discovery. An important future direction is to construct such a stochastic process by building an economic model that endogenises the decision of firms to form a cartel, the instability of collusion that would result in collapse and the discovery of a cartel by customers or a competition authority. Of particular value is learning how market conditions – such as the number firms, cost and demand variability, whether the product is an intermediate good or is sold in a retail market (so customers are either firms or consumers) – determine cartel birth, death and discovery. Such a model would provide the additional structure that could be used to more effectively extract information on the latent distribution on cartels from data on the observed distribution of discovered cartels. Progress along these lines will allow us to track the population of cartels better and eventually assess how the cartel population responds to policy initiatives and whether we are winning or losing the battle against cartels.

University of Pennsylvania
University of Southern California

Additional Supporting Information may be found in the online version of this article:

Data S1.

References

Allain, M.L., Boyer, M., Kotchoni, R. and Ponssard, J.P. (2011). 'The determination of optimal fines in cartel cases – the myth of underdeterrence', Working Paper No. 2011s-34, CIRANO.
Boyer, M. and Kotchoni R. (2012). 'How much do cartels typically overcharge?', Working Paper No. 2012s-15, CIRANO.
Bryant, P.G. and Eckard, E.W. (1991). 'Price fixing: the probability of getting caught', *Review of Economics and Statistics*, vol. 73(3), pp. 531–6.
Combe, E. and Monnier, C. (2011). 'Fines against hard core cartels in Europe: the myth of overenforcement', *Antitrust Bulletin*, vol. 56(2), pp. 235–76.
Combe, E., Monnier, C. and Legal, R. (2008). 'Cartels: the probability of getting caught in the European Union,' College of Europe, BEER Paper No. 12.
Connor, J.M. (2010). *Price-fixing Overcharges*, 2nd edn. Available at http://ssrn.com/abstract=1610262. (last accessed: 14 February 2016).
Connor, J.M. and Bolotova, Y. (2006). 'Cartel overcharges: survey and meta-analysis', *International Journal of Industrial Organization*, vol. 24(6), pp. 1009–137.
Connor, J.M. and Lande, R.H. (2012). 'Cartels as rational business strategy: crime pays', *Cardozo Law Review*, vol. 34(2), pp. 427–90.
De, O. (2010). 'Analysis of cartel duration: evidence from EC prosecuted cartels', *International Journal of the Economics of Business*, vol. 17(1), pp. 33–65.
Hammond, S.D. (2005). 'An update of the antitrust division's criminal enforcement program', ABA Section of Antitrust Law, Cartel Enforcement Roundtable, November 16.
Harrington, J.E., Jr. (2004). 'Cartel pricing dynamics in the presence of an antitrust authority,' *RAND Journal of Economics*, vol. 35(4), pp. 651–73.
Harrington, J.E., Jr. (2005). 'Optimal cartel pricing in the presence of an antitrust authority', *International Economic Review*, vol. 46(1), pp. 145–69.
Harrington, J.E., Jr. (2014a). 'Penalties and the deterrence of unlawful collusion', *Economics Letters*, vol. 124(1), pp. 33–6.
Harrington, J.E., Jr. (2014b). 'Are penalties for cartels excessive and, if they are, should we be concerned?', 13 February, at *Competition Academia*. Available at competitionacademia.com (last accessed: 14 February 2016).
Harrington, J.E., Jr. and Chang, M.H. (2009). 'Modelling the birth and death of cartels with an application to evaluating antitrust policy', *Journal of the European Economic Association*, vol. 7(6), pp. 1400–35.

Harrington, J.E., Jr. and Chang, M.H. (2015). 'When should we expect a corporate leniency program to result in fewer cartels?' *Journal of Law and Economics*, vol. 58(2), pp. 417–49.

Hinloopen, J. and Soetevent, A.R. (2008). 'Laboratory evidence on the effectiveness of corporate leniency programs', *RAND Journal of Economics*, vol. 39(2), pp. 607–16.

Hyytinen, A., Steen, F. and Toivanen, O. (2014*a*). 'Cartels uncovered', Katholieke Universiteit Leuven.

Hyytinen, A., Steen, F. and Toivanen, O. (2014*b*). 'Anatomy of cartel contracts', Katholieke Universiteit Leuven, FEB Research Report No. MSI_1231.

Ivaldi, M., Jenny, F. and Khimich, A. (2015). 'Cartel damages to the economy: an assessment for developing countries', Toulouse School of Economics.

Katsoulacos, Y. and Ulph, D. (2013). 'Antitrust penalties and the implications of empirical evidence on cartel overcharges', ECONOMIC JOURNAL, vol. 123(572), pp. F558–81.

Kawai, K. and Nakabayashi, J. (2014). 'Detecting large-scale collusion in procurement auctions', Stern School of Business, New York University.

Levenstein, M.C. and Suslow, V.Y. (2006). 'What determines cartel success?', *Journal of Economic Literature*, vol. 44(1), pp. 43–95.

Levenstein, M.C. and Suslow, V.Y. (2011). 'Breaking up is hard to do: determinants of cartel duration', *Journal of Law and Economics*, vol. 54(2), pp. 455–92.

Miller, N. (2009). 'Strategic leniency and cartel enforcement', *American Economic Review*, vol. 99(3), pp. 750–68.

Oxera Consulting LLP (2009). 'Quantitying Antitrust Damages', Study Prepared for the European Commission.

Plackett, R.L. (1965). 'A class of bivariate distributions', *Journal of American Statistical Association*, vol. 60(310), pp. 516–22.

Spagnolo, G. (2008). 'Leniency and whistleblowers in antitrust', in (P. Buccirossi, ed.), *Handbook of Antitrust Economics*, pp. 259–303, Cambridge, MA: The MIT Press.

Sweeney, B. (2006). 'The role of damages in regulating horizontal price-fixing: comparing the situation in the United States, Europe and Australia', *Melbourne University Law Review*, vol. 30(3), pp. 837–79.