DISCUSSION PAPER SERIES

No. 7761

CARTELS UNCOVERED

Ari Hyytinen, Frode Steen and Otto Toivanen

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7761.asp**

ISSN 0265-8003

CARTELS UNCOVERED

Ari Hyytinen, University of Jyväskylä
Frode Steen, NHH Bergen and CEPR
Otto Toivanen, HECER and CEPR

Discussion Paper No. 7761
March 2010

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7761

March 2010

ABSTRACT

Cartels Uncovered

How many cartels are there? The answer is important in assessing the efficiency of competition policy. We present a Hidden Markov Model that answers the question, taking into account that often we do not know whether a cartel exists in an industry or not. Our model identifies key policy parameters from data generated under different competition policy regimes and may be used with time-series or panel data. We take the model to data from a period of legal cartels - Finnish manufacturing industries 1951 - 1990. Our estimates suggest that by the end of the period, almost all industries were cartelized.

JEL Classification: L0, L4, L40, L41 and L60
Keywords: antitrust, cartel, competition, detection, hidden Markov models, illegal, legal, leniency, policy and registry

Ari Hyytinen
University of Jyväskylä
School of Business and Economics
P.O. Box 35 (MaE)
40014 University of Jyväskylä
FINLAND

Email: ari.t.hyytinen@jyu.fi

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=156617

Frode Steen
Department of Economics
Norwegian School of Economics
and Business Administration
Helleveien 30
5045 Bergen
NORWAY

Email: Frode.Steen@nhh.no

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=143159

Otto Toivanen
HECER
University of Helsinki
PO Box 17
FIN 00014 University of Helsinki
FINLAND

Email: otto.toivanen@helsinki.fi

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=145788

Submitted 19 March 2010

We would like to thank Susanna Fellman, Joe Harrington, Vesa Kanniainen, Howard Smith, John Thanassoulis and participants at the HECER conference on cartels and collusion and seminar participants in Oxford, KU Leuven and the Annual Conference of the Finnish Economic Association for comments, the Finnish Competition Authority for access to their archive, Valtteri Ahti, Janne Itkonen and Juhana Urmas for excellent research assistance and the Academy of Finland for funding. The usual caveat applies.

1 Introduction

Little is known of the efficiency of competition policy. A key reason for this state of affairs is that important statistics, such as the proportion of industries (markets) that have a cartel under an existing competition policy regime, or would have a cartel if there was no competition policy, are unknown.[1] These statistics are unknown because we lack tools to deal with incomplete cartel data. We develop a model that takes this central feature of cartel data into account and can be applied to inter-industry panel data as well as time-series data from a single industry. It allows identification of the central parameters that are relevant for competition policy, such as the probability of cartel detection by the Competition Authority (CA) and the probability of applying for leniency, if such policies are in place. We take the model to data on nationwide Finnish legal cartels in manufacturing industries from 1951 to 1990 and estimate the number of cartels in the (from a modern viewpoint counterfactual) state of no active competition policy.

To illustrate the incompletness of cartel data, consider how data on an illegal cartel is exposed. The first data point that is exposed is that the cartel exists in the period in which it is either uncovered by the Competition Authority (CA), or a member applies for leniency. The CA may then extend its investigation into the past of the cartel and eventually, either the CA and / or the court(s) establish the periods in which the cartel has existed. In reality, the cartel may have existed for longer or shorter. The CA may be able to establish that in some previous periods the cartel did not exist, or fail to establish existence in a given period either because of lack of conclusive evidence, or (as is possible in the U.S. for example), because of a plea bargain. This observation process produces data on the cartel's existence for some of the years preceding their exposure. After the investigation, a new cartel may be created in the industry, and the cycle begins again.[2] Clearly, the data available to a researcher depends on 1) the prevalence of cartels, 2) the probability of the CA exposing an existing cartel either through its investigations, or through leniency, and 3) the probability of the CA being able to establish the cartel's (non)existence in the time periods prior to exposure.

[1] The cartels we study in this paper are nation- and therefore also industrywide, covering all (e.g. regional) markets.

[2] As the cartel was abolished in the period it was discovered, the CA cannot learn how long the cartel would have existed had it not been exposed.

This data generation and exposure process, once linked to a theoretical (Markov) model of cartel behavior, maps into a Hidden Markov Model (HMM). A HMM consists of a hidden process (the industry cartel dynamics in our case) and an observation process that reveals information on the state of the hidden process for some periods, but not for others.

HMMs can be adapted to the dynamics of cartel behavior and be tailored to the specifics of the institutional environment. To show how, we build on two recent papers (Harrington and Chang 2009, Chang and Harrington 2009, HC and CH henceforth) that advance the state of the art by building a theoretical Markov model where cartels may break down. In their model, cartels face an incentive compatibility constraint (ICC) each period: If the state of the world is “too good” (demand too high, as in Rotemberg and Saloner 1986), the constraint is violated and the cartel breaks down. If there was no cartel in the previous period, the industry gets an opportunity to form a cartel with positive probability. Success in forming the cartel is subject to the ICC not being violated. HC incorporate into their model a probability of cartel detection and CH also a probability of applying for leniency, both of which lead to the cartel being exposed to the CA. The objective of these papers is to build a model where cartel births and deaths are endogenized, and to explore how data on the duration of exposed cartels could allow one to measure the effect of a policy change on the prevalence of cartels. Building on similar insights, Miller (2009) independently develops a dynamic (Markov) model of cartel formation and dissolution and studies, using aggregate data on the number of exposed U.S. cartels, whether the leniency program that the U.S. Department of Justice introduced in 1993 reduced cartellization. We map the Markov model of HC and CH into a HMM.

The modern literature on cartels can be divided into inter-industry studies and into research focusing on individual cartels in line with the New Empirical Industrial Organization (NEIO) approach surveyed e.g. by Bresnahan (1989). Prominent examples of the latter strand of the literature are Porter (1983), Lee and Porter (1984), Ellison (1994), Pesendorfer (2000), Porter and Zona (1993, 1999), Genesove and Mullin (1998, 2001), Röller and Steen (2006) and Asker (2009). These papers demonstrate the inner workings of a given cartel and as a group reveal a considerable amount of heterogeneity, both over industries and over time, in how cartels operate, how effective they are in sustaining collusive outcomes and in the welfare losses they generate.

Important precursors to our modeling approach are Porter (1983), Lee and

Porter (1984) and Ellison (1994) who all study the Joint Executive Committee, i.e., the railroad cartel for freight shipments from Chicago to the Atlantic seabord that existed in the 1880s. Porter (1983) and Lee and Porter (1984) allow for two hidden states of the industry - collusion or price-war in their set-up - and utilize an imperfect indicator to identify the actual state of the industry. Ellison (1994) extends their empirical work by bringing in a Markov structure for the hidden process. Our HMM allows for a richer state space than was necessary for these authors. Consequently, we can identify also those structural parameters of the model that describe modern competition policy (i.e., cartel detection by the CA and leniency parameters).

Prior to the emergence of NEIO, most cartel research belonged to the inter-industry approach. For example, Frass and Greer (1977) and Hay and Kelley (1974) used data on illegal cartels and found that concentration and the number of "conspirators" is positively correlated with the probability of collusion. More recently, Symeonidis' work on cartels (much of which is summarized in Symeonidis 2002) has made use of the inter-industry variation in policy changes to identify the treatment effect of cartelization. Other examples include Levenstein and Suslow's (2006b) study of international cartels, Miller's (2009) paper on the number of exposed U.S. cartels and Brenner's (2009) analysis of European Commission's leniency program.

While our data is inter-industry, our model offers the possibility of bridging these strands of the literature, as it is capable of dealing with both longitudinal inter-industry data as well as time-series data from a single industry. An inter-industry approach is necessarily less able to account for the heterogeneity in how cartels work, but it has the benefit of being able to provide answers to the policy relevant questions raised above. To illustrate how, we take our HMM model to panel data on 234 Finnish manufacturing industries from 1951 to 1990. Benefits of these data are the length of the observation period and the fact that all industries operate in the same institutional environment. In 109 of these industries, there was a known nationwide horizontal cartel in existence some time between 1951 - 1990. For the remaining 125 industries it is unknown whether a cartel ever existed, and if one did, when and for how long. We have obtained data on the 109 cartels from the Registry that the government established in 1958 when the first Finnish competition law was enacted. Such registries were commonplace in industrialized countries, including e.g. Germany, Switzerland, the Netherlands, all Nordic countries and Australia. Cartels were legal during our whole observation period. They ended in the Registry either through self-reporting or through the CA approaching them. We have gone through the Registry folder of each of these 109 cartels in detail and established those years in which one can be sure that a cartel was alive and active (i.e., the cartel engaged in correspondence with the Registry) or not in existence (dead). For some cartels, we can also establish their actual birth and/or death dates. For the remaining industry-years, we are agnostic about the (hidden) state of industries and take therefore no stance on them (i.e., such observations are assigned into state "unknown").[3] We have also collected publicly available data on industrial statistics for all 234 industries in our sample and matched them with macroeconomic variables and variables that describe the evolution of the general workings of the Finnish Cartel Registry.

We use these data to estimate the parameters of the observation process of the HMM and the associated hidden process that governed the births and deaths of the manufacturing cartels during an era when there was no active competition policy. The link to the modern era of illegal cartels is that we provide an upper bound estimate to the number of cartels - after all, while legal cartels' existence is not affected by competition policy, they are subject to (many of) the same internal incentive problems that illegal cartels face. We can therefore also answer the question: How cartelized was Finnish manufacturing in the era of legal cartels? The answer to this question is a key piece in the evaluation of modern competition policy.

By way of preview, we find that the chance of forming a cartel is around 20%, increases over our sample period and responds to positive shocks to GDP. The probability of a cartel continuing is very high (over 90%) and quite stable. Our estimate of the proportion of manufacturing industries that were in a cartel is increasing over time, and reaches more than 90% by the end of our observation period. This is true also in our counterfactual analysis where we remove the large effects of positive shocks to GDP. To the best of our knowledge, no comparable estimates exist in the literature. Our results suggest that implementation of strict competition policy is of first order importance.

The rest of the paper is organized as follows. In the next section, we first briefly review the relevant parts of the Harrington and Chang / Chang and Harrington cartel models. We then show how a HMM that matches the collusive dynamics of these models with the observed data can be specified. In the third

[3] Because of the introduction of the "unknown" state, our HMM allows us to circumvent the problem of right censoring of observed cartel durations which has plagued part of the earlier literature.

section, we describe the Finnish institutional environment vis-à-vis cartels after WWII and how the variant of the HMM that we develop takes it into account. We then move to a description of the data and key variables, present our results and discuss their policy implications. Section four concludes.

2 Modeling Cartel Births and Deaths

2.1 The Harrington and Chang model

HC and CH model an industry where (identical) firms in an industry each period simultaneously decide whether or not to collude and where collusion can be detected by a CA. From our perspective, the key difference between the papers is that the latter allows for leniency, while the former does not. Period-specific profits per firm under collusion are π; firms earn $\alpha\pi$, $\alpha \in [0, 1)$ if they compete; and a deviating firm earns $\eta\pi$, $\eta > 1$. The profit measure π has a continuously differentiable cdf H_{IC}, upper and lower bounds $\overline{\pi}$ and $\underline{\pi}$, and an expected value μ. Firms have an infinite horizon with a common discount factor δ. The CA is modelled as i) a detection (and prosecution and conviction) probability $\sigma \in [0, 1)$ and ii) penalty $F/(1 - \delta)$ paid by each firm if a cartel is exposed. CH assume that $F = \gamma(Y - \alpha\mu)$ where Y is the (scaled) continuation payoff from being in a cartel. Leniency is modeled as follows (for details, see CH):[4] Firms have an incentive to apply for leniency only if their cartel breaks down.[5] If just one firm applies for leniency, it pays a fine θF, where $\theta \in (0, 1)$, while other firms pay F. If all firms apply for leniency simultaneously, each firm pays a penalty ωF where $\omega \in (0, 1)$. It is important to note that the model does not (explicitly) allow for tacit collusion, but either the industry has a cartel that is operational, i.e., deviates from the competitive outcome (each firm earning π), or the industry is competitive (firms earning $\alpha\pi$), or a firm deviates (earning $\eta\pi$, other firms earning zero).

[4] CH model in detail the antitrust enforcement technology (see their Section 2.2). We omit this and several other parts of their modeling as they are not instrumental to our exercise.

[5] A colluding firm who is going to apply for leniency at the end of the period would be better off deviating.

At the beginning of a period, an industry is either in a cartel or not; this is dictated by the previous period's outcome. If the industry is not in a cartel, it forms a cartel with probability $\kappa \in (0,1)$. The remaining within-period sequence of events is the same for cartels thus born, and cartels that existed in the previous period: Given the realization of π (which the firms observe prior to deciding on cartel continuation), the ICC either holds, and the industry colludes, or doesn't, in which case the cartel dies. If the industry colludes, the cartel may be exposed by the CA; this happens with probability σ. It may also be exposed to the CA if at least one member of a collapsing cartel applies for leniency. In either case, the cartel is shut down and fines are levied. Finally, exposure (and death) may not happen, in which case the industry continues in state "cartel" into the next period. The structural parameters of the model are thus μ, α, η, H_{IC}, κ, δ, σ, θ, ω and F, the first six describing the industry and the last four the prevailing antitrust policy.

CH show that the ICC of an industry takes the form

$$(1-\delta)\pi + \delta[(1-\sigma)Y + \sigma(W-F)] \geq (1-\delta)\eta\alpha + \delta[W - \min\{\sigma,\theta\}F], \quad (1)$$

where Y (W) is the scaled continuation payoff from (not) being in a cartel and $F = \gamma(Y - \alpha\mu)$. Both are functions of (all of) the structural parameters. The L.H.S. of the ICC has two parts. The first denotes the current and the second the expected profits earned if there is collusion: In that case, the cartel is not exposed with probability $(1-\sigma)$ and it earns the continuation payoff is Y. With probability σ the cartel is exposed. Then the continuation payoff is W and the expected fine $F = \gamma(Y - \alpha\mu)$. On the R.H.S., the first term are the profits from deviating. Deviating will yield the competitive continuation payoff W, which is the first component of the second R.H.S. term. A deviating firm will apply for leniency if the penalty from doing so is less than the expected penalty from being caught, yielding the last component of the second term on the R.H.S. side (i.e., $\min\{\sigma,\theta\}F$).

HC set out the conditions under which cartels may be born when there is no leniency, whereas CH derive the same conditions with leniency. In both models, the expected payoff to being cartelized is defined by a recursion that can be solved through a fixed point calculation. Using the fixed point with collusion, Y^*, and rearranging (1) shows that the ICC can be rewritten in terms of π:

$$\pi \leq \phi^* \quad (2)$$

where $\phi^* = \frac{1}{(1-\delta)(\eta-1)}[\frac{\delta(1-\sigma)(1-\kappa)(1-\delta)(Y^*-\alpha\mu)}{1-\delta(1-\kappa)} - \delta(\sigma - \min\{\sigma,\theta\}\gamma(Y^* - \alpha\mu)]$ on the R.H.S is a measure of cartel stability. Cartels collapse internally if the profit shock exceeds ϕ^*. We denote the probability that this ICC is satisfied by H.

For our purposes, this modelling framework has two important features: First, it results in a Markov model for the hidden collusive dynamics of an industry and generates an unobserved sequence of cartel and non-cartel periods. Second, it incorporates the feature that some cartels are exposed because the CA detects them or because at least one member of a collapsing cartel applies for leniency.

2.2 HMM for Cartel Births and Deaths

HMMs provide a means to study dynamic processes that are observed with noise.[6] The evolution of the population of cartels matches this description, because we typically observe the (collusive) dynamics of an industry only irregularly, if at all, and only for discovered cartels. A HMM consists of an underlying hidden ("unobserved") process and an observation process. In particular, the observed data, O_{it} , for industry $i = 1, ..., N$ and periods $t = 1, ..., T_i$ follow a HMM if the hidden states, $\{Z_{it}\}_{t=1}^{T_i}$, follow a Markov chain and if, given Z_{it}, observation O_{it} at time t for i is independent of the past and future of hidden states and observations (see Appendix A for a more detailed description). In our case, the hidden process is the state of the industry and the observation process is what the researcher knows about the state of the industry in a given period. The dimension of the state space of the hidden process is typically either assumed or estimated. In our case, it follows directly from economic theory and the institutional environment.

[6] Our model belongs to the class of finite Hidden Markov Models (e.g., Cappé, Moulines and Rydén 2005, pp. 6). HMMs are not very widely used in economics. There have been some applications in macroeconomics (and time-series analysis), where they are known as Markov switching models (see, e.g., Hamilton 1989, and Engle and Hamilton 1990). Other examples include Bhar and Hamori (2004) and Wang and Alba (2006).

2.2.1 Hidden Process for Cartel Births and Deaths

Consider cartel births and deaths in some industry i at time $t > 1$. At the beginning of period t the industry is either in a cartel or not. Being in a cartel is here synonymous with actually colluding (i.e., the market outcome not being competitive).[7] If the industry is not in a cartel, a cartel is formed with probability κ_{it}. Conditional on the opportunity, the cartel is stable and becomes operational with probability H_{it}. This probability is directly linked to the ICC (2) discussed above. If the industry is in a cartel at the beginning of period t, then it stays alive with probability H_{it}. With probability $1-H_{it}$, an existing cartel breaks down during period t.

There is a CA that constantly monitors the status of each industry. At the end of period t, the state of industry i is detected by the CA with probability σ_{it}. If the industry is in a cartel, the cartel is shut down immediately (and fines are levied). If the industry is not in a cartel, the industry stays as is. Besides the CA, there is a corporate leniency program in place. Conditional on the cartel breaking up, the probability that it will be exposed to the CA because of a leniency application is ν_{it}.

This process for cartel births and deaths means that at the end of period t, industry i is either not in a cartel ("n"), is in an on-going cartel ("c"), has been detected and shut down by the CA ("d") or has after the break up been exposed to the CA because of a leniency application ("l"). Treating these four outcomes as the states of hidden process for Z_{it}, its state space is $S_Z = (n, c, d, l)$. The associated transition matrix $\mathbf{A}_{it}$ is

$$\begin{bmatrix} (1-\kappa_{it}) + \kappa_{it}(1-H_{it})(1-\nu_{it}) & \kappa_{it}H_{it}(1-\sigma_{it}) & \kappa_{it}H_{it}\sigma_{it} & \kappa_{it}(1-H_{it})\nu_{it} \\ (1-H_{it})(1-\nu_{it}) & H(1-\sigma_{it}) & H_{it}\sigma_{it} & (1-H_{it})\nu_{it} \\ (1-\kappa_{it}) + \kappa_{it}(1-H_{it})(1-\nu_{it}) & \kappa_{it}H_{it}(1-\sigma_{it}) & \kappa_{it}H_{it}\sigma_{it} & \kappa_{it}(1-H_{it})\nu_{it} \\ (1-\kappa_{it}) + \kappa_{it}(1-H_{it})(1-\nu_{it}) & \kappa_{it}H_{it}(1-\sigma_{it}) & \kappa_{it}H_{it}\sigma_{it} & \kappa_{it}(1-H_{it})\nu_{it} \end{bmatrix}. \tag{3}$$

Besides the fact that the rows of the matrix sum to one, there are three features that we should emphasize about the transition matrix:

[7] In other words, a "price war" would be classified as a period of no cartel. This is in line with the HC model where the deviation by one firm leads to the cartel collapsing and the industry entering the next period in state "no cartel".

First, the elements of the matrix are the transition probabilities of a first-oder Markov chain. The cell in the upper left-corner, for example, gives $P(Z_{it} = n|Z_{i,t-1} = n)$=$(1-\kappa_{it}) + \kappa_{it}(1-H_{it})(1-\nu_{it})$ =$1+\kappa_{it}\,(H_{it}(1-\nu_{it})+\nu_{it})$. It is derived as follows: If an industry is not in a cartel at $t-1$, then with probability $(1-\kappa_{it})$ there is no opportunity to form a cartel. If there is an opportunity, the newly born cartel may turn out to be unstable, but the member firms do not apply for leniency. The probability of this event is $\kappa_{it}(1-H_{it})(1-\nu_{it})$. The probability given in the upper left-corner cell is the sum of the probabilities of these two events.

Second, the detection probability σ_{it} shows up only in columns 2 and 3 because the detection activities of the CA affect only those states in which an industry is in a cartel at the beginning of period t. The cell in the first row of the third column, for example, gives the probability for the event that an industry that has not been in a cartel at $t-1$ forms a cartel during period t but is immediately detected and shut down by the CA.

Third, the first and two last rows are equal, because we assume that if an industry has at $t-1$ been in a cartel that has been exposed to the CA, it does not affect the process that leads to the creation of new cartels in subsequent periods. This assumption could be relaxed at the cost of additional complexity.

To complete our specification of the hidden process for cartel births and deaths, let the R.H.S. of (2) vary over industries and time and rewrite the inequality by substracting from both sides the mean of the expected profits under collusion in industry i during period t (μ_{it}). This leaves a demeaned profit shock, $\pi_{it}-\mu_{it}$, to the L.H.S. of the inequality, now taking the form of a discrete choice equation with a particular structure on the R.H.S. With H_{it} denoting the probability that the inequality holds for industry i in period t, we have

$$H_{it} = H_{ICDM}(\phi^*_{it} - \mu_{it}) \tag{4}$$

where $H_{ICDM}(\bullet)$ refers to the c.d.f. of the demeaned profit shock. We can think of $\phi^*_{it}-\mu_{it}$ as a function of observable characteristics (which could enter, e.g., through μ_{it}) and the structural parameters of the model.

2.2.2 Observed Data and the Observation Process

In modern data sets on discovered cartels (see, e.g., Miller 2009, Brenner 2009, Levenstein Suslow 2006a,b), the observed data, O_{it}, for industry i at time t vary but are becoming increasingly detailed. For some industries in these data sets, the state of the industry can be determined for certain periods from the records and publications of the CA and the courts. For some (but clearly not for all) industries, these records allow one to infer, for example, whether industry i was or was not in an on-going (operational) cartel in a given period. Specifically, the records may show whether the industry has been subjected to a competition policy action (i.e., whether a cartel has been detected by the CA or whether it was exposed because some, perhaps unidentified members applied for leniency). For the remaining industry-period observations in the data, the state of the industry cannot be determined at all, or perhaps not with any acceptable level of confidence. In particular, for a number of industries, the status of the industry cannot be determined for any period. A prime example of such a case is an industry that has never been investigated or convicted for having a cartel.[8]

There are three important ways in which the records and publications of the CA and courts may be incomplete. First, information about the status of an industry *beyond* the detection is typically available only for the cartels detected by the CA or for those cartels that were exposed to the authorities because of a leniency application. Second, even for the exposed cartels, the status of the industry may be determined for some periods only. Third, the (published) records of the authorities do not cover all industries and thus cannot provide information about the unrecorded industries' states.

The foregoing suggests that in each period t, either the state of industry i is not known ("u"), or the industry is observed not to be in a cartel ("n"), to be in an on-going cartel ("c"), to have been been detected and shut down by the CA ("d") or to have been exposed to the CA because of a leniency application ("l"). These five observed cartel outcomes give the state space of the observation process, $S_O = (n, c, d, l, u)$.

HMMs model such observed data by linking it to the hidden process that govern the formation and dissolution of cartels. When the unobserved state of industry i at time t is $k \in S_Z = (n, c, d, l)$, a HMM postulates that the probability of observing $w \in S_O = (n, c, d, l, u)$ is

[8]See Appendix B for a clarifying example.

$$b_{it}^k(w) = P(O_{it} = w | Z_{it} = k) \quad (5)$$

To derive the observation probabilities explicitly and to match them to a modern competition policy environment, we make the following assumptions:

First, we assume that if an industry is not in a cartel, its (true) state is observed in the data available to the reseacher with probability $b_{it}^n(n) = \beta_{it}^n$. If this event happens, $O_{it} = Z_{it} = n$. With the complementary probability $b_{it}^n(u) = 1 - \beta_{it}^n$, the state cannot be determined realiably and remains unknown. If this event happens, $O_{it} = u$ and $Z_{it} = n$. If an industry is in a cartel, its (true) state is observed in the data with probability $b_{it}^c(c) = \beta_{it}^c$. In this case, $O_{it} = Z_{it} = c$. Again, with the complementary probability, the status remains unknown. This implies that $O_{it} = u$ and $Z_{it} = c$.

This formulation of the observation process relies on the assumption that if an industry is (is not) in a cartel, the observed data never wrongly suggest that it is not (is). This assumption imposes $b_{it}^n(c) = b_{it}^c(n) = 0$ and implies that there are no mistakes in the records and publications of the CA and the courts. We stress that this restriction may sound stronger than it is, because if and when one has reasons to suspect that there are such errors, the status of an industry can be labelled "unknown". Moreover, this assumption can be relaxed if the data contain information about potential mistakes or mislabelings in the records.[9]

The second assumption that we make to derive the observation probabilities is that the exposure of a cartel to the CA is observed (by the researcher) with probability one. Formally, we impose $b_{it}^k(n) = b_{it}^k(c) = b_{it}^k(u) = 0$ for $k = d, l$ together with $b_{it}^d(d) = b_{it}^l(l) = 1$. We think that it also is plausible to assume that the observed data never suggest (to the researcher) that a cartel has been shut down by the CA or exposed because of leniency when it really was not (i.e., $b_{it}^n(d) = b_{it}^c(d) = b_{it}^l(d) = 0$ and $b_{it}^n(l) = b_{it}^c(l) = b_{it}^d(l) = 0$). [10]

The estimates of β_c^{it} and β_n^{it} reflect the ability of the CA (and courts) to determine, in an ex post investigation, whether a detected cartel did or did not exist in the periods prior to the detection. They are therefore potentially policy

[9] Porter (1983), Lee and Porter (1984) and Ellison (1994) allow for mistakes in the observation process, but do not have the observation state "unknown".

[10] If, for some reason, there is uncertainty about the cause of exposure, the model can be extended to allow for such an observation process. One way to do so is to let $b_{it}^l(l) = \beta_{it}^l$ with $b_{it}^l(d) = 1 - \beta_{it}^l$ and $b_{it}^d(d) = \beta_{it}^d$ with $b_{it}^d(l) = 1 - \beta_{it}^d$.

relevant.

2.2.3 HMM Representation

Assuming an initial distribution for Z_{i1} (i.e. the probability that unit i is at the unobserved state $k \in S_Z$ in the initial period),

$$\tau_i^k = P\left(Z_{i1} = k\right) \tag{6}$$

and combining the hidden process for cartel births and deaths with the observation process results in the following HMM representation for industry i:

$$\begin{aligned} \mathbf{Z}_i &= Z_{i1}, Z_{i2}, ..., Z_{iT_i} \\ \mathbf{O}_i &= O_{i1}, O_{i2}, ..., O_{iT_i} \\ S_Z &= (n, c, d, l) \\ S_O &= (n, c, d, l, u) \end{aligned} \tag{7}$$

where $\mathbf{Z}_i$ is the vector of the actual states of industry i that take values from the state space S_Z and $\mathbf{O}_i$ is the vector of observed states for industry i that take values from the observation state space S_O. The remaining elements of the HMM are transition probabilities $\mathbf{A}_{it} = \left[\begin{array}{c} a_{it}^{jk} \end{array}\right]$ given in equation (3) and observation probabilities

$$\mathbf{B}_{it} = \left[\begin{array}{c} b_{it}^k(w) \end{array}\right] = \left[\begin{array}{ccccc} \beta_{it}^n & 0 & 0 & 0 & 1-\beta_{it}^n \\ 0 & \beta_{it}^c & 0 & 0 & 1-\beta_{it}^c \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]. \tag{8}$$

To derive the likelihood of the HMM, let Θ denote the model parameters, $\mathbf{D}_{i1}$ a (5×1) vector with elements $d_{i1}^k(w) = \tau_i^k b_{i1}^k(w)$, $\mathbf{D}_{it}$ a (5×5) matrix with elements $d_{it}^{jk}(w) = a_{it}^{jk} b_{it}^k(w)$ for $t > 1$, and $\mathbf{1}$ a (5×1) vector of ones. As shown in e.g. MacDonald and Zucchini (2009, p. 37), Altman (2007) and Zucchini, Raubenheimer and MacDonald (2008), the likelihood for the whole observed data can be written as

$$L(\Theta; \mathbf{o}) = \prod_{i=1}^{N} \left\{ \left(\mathbf{D}_{i1}\right)' \left(\prod_{t=2}^{T_i} \mathbf{D}_{it} \right) \mathbf{1} \right\} \tag{9}$$

where $\mathbf{o}$ denotes the data (the realization of $\mathbf{O}$).[11] To maximize $L(\Theta; \mathbf{o})$, (direct) numerical maximization methods can be used (MacDonald and Zucchini 2009, Chapter 3; Turner 2008). Typically, a normalization (scaling) is used to avoid numerical underflow.[12]

The HMM summarized above can be extended to allow for unobserved heterogeneity (see e.g. Altman 2007). To bring in unobserved heterogeneity properly into our HMM would however require modeling it within the theoretical model. We proceed without allowing for unobserved heterogeneity.

2.2.4 Estimation

Because $\{\tau_i^n, \kappa_{it}, H_{it}, \sigma_{it}, \nu_{it}, \beta_{it}^n, \beta_{it}^c\}$ are all probabilities, a simple way to parametrize them is to assume a standard probability model for each of them. This can mean, for example, that one imposes $\kappa_{it}= \Phi\left(\kappa' \mathbf{x}_{it}\right)$, where $\Phi(\bullet)$ is the c.d.f. of the normal distribution, $\mathbf{x}_{it}$ denotes the available explanatory variables and κ is the parameter vector to be estimated. We adopt this appproach in our empirical application.

The identification of (the parameters of) a general finite HMM follows from the identifiability of mixture densities (see Cappé, Moulines and Rydén 2005, pp. 450-457). The parameters of our HMM for cartel births and deaths are identified for further two reasons: First, we impose theoretical restrictions upon the hidden process describing the formation and dissolution of cartels. These restrictions allow us to circumvent the problem of identifying the dimension of the hidden process: The theoretical model and the institutional environment (e.g. the existence of a corporate leniency facility) directly suggest that $S_O = (n, c, d, l)$, which implies that the dimension is four.[13] A second source of identification are

[11] Picking the appropriate elements from $\mathbf{A}_{it}$ and $\mathbf{B}_{it}$, we can determine $d_{it}^{jk}(w) = a_{it}^{jk} b_{it}^{k}(w)$ for $t > 1$, i.e., the elements of matrix $\mathbf{D}_{it}$ of the likelihood function that is given as equation (9). If, for example, $o_{it} = c$, the upper left-corner cell of $\mathbf{D}_{it}$ is $d_{it}^{nn}(w) = a_{it}^{nn} b_{it}^{n}(c) = 0$. For $t = 1$, the elements of the vector $\mathbf{D}_{i1}$, $d_{it}^{k} = \tau_i^k b_{i1}^k(w)$, in the likelihood function can be determined similarly.

[12] It is a well-known property of (9) that even for a short time period, the individual elements of the vector of probabilities (held at a given stage of the maximization algorithm) often become indistinguishable from zero.

[13] For a general discussion of identification, see Cappe Moulines, and Rydén (2005, ch. 15).

the parameter restrictions that we impose on $\mathbf{B}_{it}$.

An intuitive way to think about the identification of our HMM is that we have only 4+2 probabilities that call for identification, but a greater number of moments (transitions) that identify them. The observed transitions from c to c and c to n identify H_{it}, whereas the observed transitions from n to c and n to n identify κ_{it}. Transitions from and to d and from and to l allows us to identify σ_{it} and ν_{it}. Finally, differences in the transition rates from c to c and c to n and from n to n and n to c identify β_{it}^{n} and β_{it}^{c}.

Estimation of the parameters of H_{it}, as given by (4), deserves a further comment. One way to proceed is to estimate a reduced form of this probability. This is computationally simpler, but has the downside of not enabling a counterfactual analysis of different competition policy regimes. The reason for this is that the effects of the competition policy parameters σ_{it} and ν_{it} on H_{it} are not identified. The other possibility is to estimate H_{it} structurally, but this requires that the fixed point with collusion (Y^*) and the associated threshold (ϕ^*) are computed. The estimation routine could be e.g. a nested fixed point algorithm where one starts from some initial values for the estimated parameters, calculates the fixed point (i.e., the value of ϕ^*), proceeds to re-estimate the structural parameters by Maximum Likelihood, and continues until convergence is achieved.[14] Alternatively, the recently introduced MPEC algorithm could be utilized (Judd and Su 2008).

2.2.5 State Prediction

A convenient feature of HMMs is that the hidden states of the underlying Markov model can be analyzed in a relatively straightforward way (see Appendix C for a more detailed description of some of these methods). For us, it suffices to note that the HMM allows for period-by-period inference about the state of the Markov chain that is most likely to have given rise to the observed data for a given industry in a given period t. This procedure is called 'local

[14] Natural candidates for initial values would be the parameter estimates from a model where H_{it} has been modelled in reduced form. An issue one would have to solve is how to deal with the potential multiplicity of Y^*. See CH for a discussion of multiple equilibria in their model.

decoding'. In a cartel application, this feature means that one can, for example, deduce the likelihood for the existence of a cartel in a given industry for those periods for which the observed data are not directly informative about the state of that industry. We will apply this method to answer the question of how the cartelization of the Finnish manufacturing industries evolved within our observation period.

3 Finnish Legal Cartels

3.1 Modelling Finnish Legal Cartels

3.1.1 The Institutional Environment and the Cartel Registry

The Finnish institutional environment vis-á-vis cartels mirrors wider European and especially Swedish developments both before and after WWII. Before the war, and until 1958, there was no competition law. The apparent reason was that the then prevailing liberal view held that contractual freedom entailed also the right to form cartels (see Fellman 2009). We have chosen to focus on the developments after 1950, because the Finnish economy was heavily regulated during WWII and because those regulations were mostly lifted by 1950.

Finnish steps towards a legal framework for competition policy started soon after WWII as in 1948, a government committee was set to provide a framework for competition legislation. The ensuing legislation (the 1958 law) was built around the idea of making cartels public through registration. Registration, however, was to be done solely on authorities' request. Only tender (procurement) cartels became illegal, and even these were apparently not effectively barred from operation (Purasjoki and Jokinen 2001). Vertical price fixing could be banned if deemed "particularly harmful". The law embodied the prevailing thinking of cartels not (necessarily) being harmful. A Finnish CA was set up to register the cartels. Here Finland followed Norway and Sweden, which set up similar registers in 1926 and 1946.

The CA was active, sending out 9750 enquiries by 1962, and registering 243 cartels (Fellman 2009, pp. 17). However, the fact that registration was dependent on authorities' activism was an issue. To tackle this, the law was revised and the new law took effect in 1964. The main new feature was compulsory notification by those cartels that established formal bodies such as associations.

Cartels purely based on agreements between firms without formal organizations were still exempt from compulsory registration. The law was again revised in 1973. The single largest change appears to have been that the obligation to register was again widened. Finland finally edged towards modern competition law with a committee that started its work in 1985. On the basis of this work, a new law took effect in 1988. This law gave the newly established Finnish Competition Authority (new FCA) the right to abolish agreements that were deemed harmful. The law also abolished cartel agreements' status as legally binding contracts. The new FCA initiated a negotiation round with cartels where these were asked to provide reasons why they should be allowed to continue. In 1992 the law was again changed (and took effect Januar 1^{st} 1993): Now cartels became illegal. Some exceptions were however allowed.

The former and current Director Generals of the Finnish CA (Purasjoki and Jokinen, 2001) sum up the environment prior to the 1988 law as follows: "Time was such that there seemed no need to intervene even in clear-cut cases, especially if they had been registered. Registration had been transformed into a sign of acceptability of the [cartel] agreement, at least for the parties involved [in the cartel]". Based on this, we end our analysis to 1990.

For each cartel in the Registry, there is a folder containing the entire correspondence between the Registry and the cartel (members). For many cartels, the cartel contract is also available (as it was requested by the Registry). This correspondence allows one in most cases to establish e.g. who took the initiative that led to a registration, or a change in it. As we explain below, this information allows one to pin down the actual birth and/or death dates of some cartels and/or their existence (or non-existence) in certain industries and years.

Over the period of its existence the Finnish Cartel Registry registered as many as 900 cartels. For each cartel, the Registry assigned a 4-digit ISIC code. We concentrate on manufacturing industries and particularly on the 109 industry-wide manufacturing cartels that are found in the Registry. The total number of ISIC-4 manufacturing industries is 234, and we follow these industries over 40 years from 1951 to 1990.

3.1.2 HMM for Legal Cartels

The institutional context that has generated our cartel data can be modelled using a nested version of the HMM for cartel births and deaths developed above.

Due to cartels being legal, the Finnish CA did *not* attempt to close cartels over the period 1951-1990, nor was there a leniency program in place. These features of the data can be captured by imposing $\sigma_{it} = 0$ and $\nu_{it} = 0$ for all t and i. This restriction reduces the state space of the hidden process to $S_Z = (n, c)$ and the dimension of the observation process to $S_O = (n, c, u)$, because the state of an industry can never be d (i.e., no cartel was ever shut down by the Finnish CA) or l (if a cartel broke down, it never resulted in a leniency application).

Taking into account these modifications, noting that the panel of Finnish legal cartels is balanced and assuming constant initial probabilities of having a cartel ($\tau_i^k = \tau^k$ for $t = 1$), results in a "legal-era" HMM with $N = 234$, $T = 40$ and with transition probabilities

$$\mathbf{A}_{it} = \begin{bmatrix} a_{it}^{nn} & a_{it}^{nc} \\ a_{it}^{cn} & a_{it}^{cc} \end{bmatrix} = \begin{bmatrix} (1 - \kappa_{it} H_{it}) & \kappa_{it} H_{it} \\ (1 - H_{it}) & H_{it} \end{bmatrix} \tag{10}$$

and observation probabilities

$$\mathbf{B}_{it} = \begin{bmatrix} b_{it}^{n}(n) & b_{it}^{n}(c) & b_{it}^{n}(u) \\ b_{it}^{c}(n) & b_{it}^{c}(c) & b_{it}^{c}(u) \end{bmatrix} = \begin{bmatrix} \beta_{it}^{n} & 0 & 1 - \beta_{it}^{n} \\ 0 & \beta_{it}^{c} & 1 - \beta_{it}^{c} \end{bmatrix}. \tag{11}$$

Despite the reduction of the state spaces, the interpretation of the HMM and its various components does not change, except that H_{it} no longer depends on σ_{it} or ν_{it}. The interpretation of β_{it}^n and β_{it}^c nevertheless deserves a comment, as they now reflect the ability of the Finnish CA to identify and keep record of cartelized industries during the era when cartels were legal in Finland. Because $\beta_{it}^n \leq 1$ and $\beta_{it}^c \leq 1$, the model explictly allows for the possibility that there are holes in the Registry and thus that the state is unknown for some industry-years in the data. There are two primary reasons for this incompleteness: First, information about the state of a registered cartel can be incomplete over time. It is, for example, possible that the cartel did not reply to the inquires by the CA about its status. On the other hand, some cartelized industries were never registered and some industries may not have had cartels.[15] For these cases, our data conservatively assign state u, as we explain in greater detail below.

[15] Purasjoki and Jokinen (2001) mention a few cartels that were not registered, but they do not explain how these cartels were exposed (apart from them being exposed as part of the negotiation initiative set up by the new FCA in the late 1980s). This neverteless confirms that the Registry was not complete.

3.2 Data Sources and Description

Our data come from two main sources, Statistics Finland and the Finnish Cartel Registry. The former provides us with macroeconomic data and 2-digit ISIC level industrial statistics that we match to the 234 4-digit manufacturing industries in our sample.[16] The latter allows us to identify 109 nationwide manufacturing cartels and is our sole source of cartel data.

3.2.1 The Definition of States and Observed Transitions

For each of the 109 manufacturing cartels in the Registry, we have been through their entire folder using a "semi-structured" approach.[17] For all these cartels we know when they entered and exited the Registry. Several cartels were born prior to the Registry being established, and some cartels did not register immediately at the time of birth (of either the cartel or the Registry), but some years later. Correspondingly, we have cartels that died but whose death was not found out until some years after their death. To improve our data we therefore used additional Registry information on these cartels where available to establish actual birth and death dates. The Registry also allows us to date changes made to cartel contracts or in the organization of cartels. These events yield information on the cartel being alive. We use all this information to define the states of the observation process for a given industry in a given year.

To be more precise, the Registry contains information on seven types of events that the registered cartels (may) have experienced between 1951-1990. First, we know for all the cartels when they entered the Registry ('register birth'

[16] We use 2-digit ISIC data because of difficulties in tracking industries across three changes in the 4-digit industry definitions that take place during our observation period. As the data was not available in electronic form, we collected data for every 4th years and interpolated the values in between.

[17] After initial discussions on what it is that we want to record, we randomly chose 8 cartels and had 4 researchers (including two of us) go independently through the material to establish whether the information we sought was available, and if, how to record it. We then checked the 4 individuals' records against each other, and decided on a common approach and interpretation of e.g. various wordings that we encountered. We then followed a written protocol in collecting the information.

- t^{rb}). In addition, we know the date when they exited the Registry for 92 cartels ('register death' - t^{rd}). The Registry also has information on different occasions when the cartels made changes in their cartel contracts ('contract change' - t^{cc}), such as an addition or a deletion of members. There can be many events like that per cartel over the years. From the additional archive information that we were able to collect, 71 cartels' actual birth ('birth' - t^b) can be established, and for 76 cartels we know the actual death date ('death' - t^d). In addition, there were incidences where a cartel was observed to be operational prior to the registered birth ('actually alive' - t^{aa}) and also some incidences where we found proof of the cartel being still alive after their registered birth and before their (registered) death ('still alive' - t^{sa}).

We use these seven types of events to define what the observed state of industry i was in year t. The observation state space is $S_O = (n, c, u)$ and we assign all industry-year observations into one of these three states. How we do this is illustrated in Figure 1. It is important to keep in mind that our interpretation of state c is (in line with HC) that not only was there a cartel agreement in place, but also that the cartel was active in the sense that the firms were trying to avoid the fully competitive market outcome. Similarly, state n is interpreted to mean that the industry was competing. Any observations for which we cannot give such an interpretation are assigned into state u. Importantly, this means that if an industry does not show up in the Registry at all, all observations for it are assigned into u.

[Figure 1 – Timeline for state-definition and observed cartel incidences here]

Cartels for whom we observe the actual birth date t^b or for whom we have information on the cartel being actually alive some year prior to register birth (t^{aa}) are assumed to be alive between t^b (t^{aa}) and the date of register birth (t^{rb}). Correspondingly, cartels for whom we know the actual death date (t^d) are presumed to be dead between t^d and the date of register death (t^{rd}). In addition, cartels are assumed to be alive every year where we either observe a 'still alive' incidence or a 'contract change' incidence. Cartels are also assumed to be alive the year before their actual death, when such can be established. Finally, cartels are assumed dead the period prior to actual birth. For all the other periods, the state of the observation process is u (unobserved).

The definition of the observed states is in our view quite conservative. For instance, even though the Registry effectively assumed that the cartels were alive between t^{rb} and t^{d} , we only assign an industry into state c when an event like t^{sa} or t^{cc} appears. The reason for including the period between t^{b}/t^{aa} and t^{rb} as observed c-states is due to the assumption that when a cartel is asked to register (at t^{rb}), it had no reason to tell any other birth date but the latest. Correspondingly, when the Registry finds out that the cartel is dead (t^{rd}), there is no incentive for the cartel not to inform the Registry of an actual restart between t^{rb} and t^{rd} when they are confirming their death to the Registry. This motivates recording the periods between these incidences as state n. Note also that the way in which we define observed/unobserved states here removes the usual problem of right censoring for cartels where we do not know the ending date, since all cartels that apparently lived when the Registry ended in 1990 are coded as u after their last t^{sa}/t^{cc} incidence.

Based on this coding we end up with 939 (industry-year) observations where we know the actual status of the cartel. For 365 of these, we know the industry is not in a cartel (n-states) and for 574 we can establish the existence of a cartel (c-states). For the remaining 8421 observations the status of the industries is unobserved. Thus in the sample where all industries are included, we assing n or c for 10% of the observations. If we disregard the 125 industries where no cartels are registered, we have 5000 (125×40) fewer u observations.

3.2.2 Data Description

In the prior literature, register data are often assumed to be roughly in line with the underlying true distribution of cartel births and deaths. Clearly this is not the case in our data. During the first few years of the Registry, several cartels that were born earlier entered the Registry. Correspondingly, exit from the Registry is very often later than the actual death. If we look at the average differences, the representative cartel was born 3.6 years earlier than it was registered and died 2.6 years earlier than it exited the Registry. This suggests that cartel duration is biased downward when using entries into and exits from cartel registers as indicators of cartel durability.

From the Registry data we observe few cartels lasting more than 20 years; cartels lasting between 4-6 and 16-20 years are the most common. We also find that if the dates of entries into and exits from the Registry were used, we would find too few short lived cartels (1-3 years) due to late registration of cartel deaths. The real numbers suggest that the modal cartel lives for 4-6 years.

Levenstein and Suslow's (2006b, Figure 1) analysis of 72 illegal international cartels reveals also that the most common duration of cartels is 4-6 years. It is interesting that our legal cartel survival distribution (real) is in line with what they found for illegal cartels.

[Figure 2 – Density on real survival, Registry survival here]

Turning to the estimated density of register-based duration, displayed in Figure 2, we find it to be skewed to the left, with more of the probability mass in the interval 5-25 years. Furthermore, the density displays two clear humps. Looking at the density of the actual duration we find that the hump in the density around years 16-20 disappears. The median for the real duration is 10 years, whereas the median based on register durations is 11 years. The averages mirror the overall skewness with an average duration of 13 (12) years using the real (Registry) numbers.

Our cartels are thus of somewhat longer duration than what others have found (Levenstein and Suslow, 2006a, Table 1 and 2006b). However, most of these papers are studying illegal cartels. The closest study to ours is Jacquemin, Nambu and Dewez (1981) who, studying legal Japanese export cartels, find an average duration of 10 years.[18]

The variation in cartel status over time is shown in Figure 3. Typically we have more cartel observations (c-states) during the first 15 years of the Registry's existence, with a peak in 1959. In this period we have few "no cartel" observations (n-states). During the early eighties the annual share of n observations is double the share of c observations.

[Figure 3 – n,c,u development here]

[Table 1 – Observed transitions here]

[18] It is also interesting to look at how the duration of cartels changes over the sample period. For the cartels that were born prior to 1966 (half of the subsample of 52) the average real duration was 18.8 years as compared to a Registry duration of 15.6. Cartels that were born later than this had average durations of 7.2 and 8.6 years.

In Table 1 we show the transitions from period $t-1$ to period t that follow from our definitions of the three states. The difference between considering the cartelized industries only and all the industries is that in the latter case, we observe a lot more transitions from u to u. For those industries with a registered cartel, 78% of the observations are transitions from u to u whereas in the whole data, the proportion is 90%. Adding the industries that do not have an exposed cartel obviously yield no more information on transitions from state n to c or vice versa, but crucially, do affect the cell probabilities.

3.2.3 Explanatory Variables

Workings of the Registry

It seems plausible that the ability of the Registry to detect the births and deaths of cartels improved or at least varied over time. To accommodate this, we make the two observation probabilities (β_{it}^{c} and β_{it}^{n}) each a function of two variables: First, we let β_{it}^{c} (β_{it}^{n}) vary with the number of cartels that entered (exited) the Registry in year $t-2$. Second, we allow β_{it}^{c} (β_{it}^{n}) be a function of the (once) lagged cumulative number of registered births (deaths). These variables are denoted ($Birth-flow$, $Birth-stock, Death-flow, Death-stock$) and they are computed using the data from the whole Registry with 900 cartels. Their time series variation is shown in Figure 4.

[Figure 4 – Birth and death flow and stock here]

It is interesting to note that even though there is a weak negative trend in the number of annually registered cartels, there is also a lot of variation over time. There is an upward trend in the number of Registry deaths.

Macroeconomic Demand Fluctuations

There is a large cartel literature focusing on the importance of demand and business cycle fluctuations for cartels. Most notable are Green and Porter (1984), whose model suggests that price wars will arise in response to unobserved negative demand shocks, and Rotemberg and Saloner (1986), whose model predicts price wars during booms (later discussed by e.g. Haltiwanger and Harrington 1991).[19] Levenstein and Suslow (2006a) summarize this literature and suggest that it is not obvious how cartel survival relates to fluctuations in demand. Of the more than 50 studies and 19 cartels summarized in their Table 1, 6 were formed during downturns, whereas 10 were not. The literature suggests that cartel formation may be linked to the growth trend as well as to idiosyncratic changes in demand not anticipated by the cartel.[20]

We have a long panel with 40 years of data over a period in which the Finnish macroeconomy went through large changes. This makes our dataset particularly useful for studying whether cartel behavior is affected by business cycle changes. To account for this, we include several macroeconomic variables into the HMM. We detrend the Finnish GDP over the 40 years period using the Hodrick and Prescott filter (Hodrick and Prescott, 1997), decomposing GDP into the long run growth trend ($HP-trend$) and deviations from the long run trend.[21] We follow Levenstein and Suslow (2006b) and decompose the deviations into two variables, one capturing positive deviations from the long run trend ($GDP-pos$), and one capturing all negative deviations from the long run trend ($GDP-neg$).[22] How these three variables and the Finnish GDP evolve over our sample period can be seen from Figure 5. The positive GDP shocks in e.g. early 1970s and late 1980s are especially worth pointing out.

[Figure 5 – Graph of GDP, HP-trend, GDP-neg, GDP-pos here]

[19] Of course, in these models the cartel is shown to exist under certain conditions, and the price wars are part of the equilibrium path of the cartel. It is not clear how to relate periods of active collusion and periods of punishment to the births and deaths of cartels.

[20] Some examples of empirical studies include Jacquemin, Nambu and Dewez (1981), Suslow (2005) and Levenstein and Suslow (2006b).

[21] We apply annual GDP volume data and generate the HP-trend and the deviation using a smoothing index of 100.

[22] Note that both variables are defined in absolute terms.

Industry Characteristics

Several authors have focused on the importance of industry characteristics when explaining cartel formation (see Levenstein and Suslow 2006a, for a survey). Slade (1989, 1990) suggests, for example, that price wars can arise from changes in industry characteristics. Cartel members' knowledge of fundamental structural parameters may be incomplete, and industry specific shocks (e.g., negative sales shocks) will change the equilibrium prices. We therefore include the gross value of production over time (GVP), as measured at the level of 2-digit industries. Even though the gross value of production is correlated with GDP over time at the macroeconomic level, there is large of variation across industries.

Among others, Bradburd and Over (1982) argue that organizational costs of both cartel formation and maintenance are expected to increase with the number of firms in an industry. We do not have an ideal measure for the number, but can nevertheless include the number of plants, as measured at 2-digit level ($Plants$). There is a lot of variation in this variable across industries, but in the aggregate there is no clear trend in it over time.

We include the ratio of raw material expenses to the gross value of production ($Materialshare$) as a measure of (average) variable costs of production. The ratio of blue collar working hours to the gross value of production ($Hours/GVP$) is a measure of (the inverse of) labor productivity. All our covariates are summarized in Table 2.

[Table 2 - Descriptive statistics here]

3.3 Estimation Results

We parameterize the probabilities of our legal era HMM as single index functions of the explanatory variables described above, using the c.d.f. of the standard normal distribution. Given that the policy parameters that enter the ICC are zero in our data (due to cartels being legal), the gain from estimating H_{it} structurally is very minor. We therefore estimate a reduced form of it. The HMM is estimated with ML, using the likelihood function (9).

We present two sets of estimation results in Tables 3-5: In the first set, we fix the initial probability (the probability of being in a cartel in period $t = 1$)

to 0.5; in the second, we estimate it. In column 1 we report the results from a model without covariates (Model 1, M1); in column 2 those from a model with only macro variables (Model 2, M2); and in column 3 results from a model where industry characteristics have been added (Model 3, M3).

[Table 3-5 - Estimation results here]

Starting from a specification test, we can reject the null hypotheses that the coefficients of the covariates are jointly zero in models that use fixed and estimated τ^n. Focusing then on H_{it} (Table 3), we find the trend in GDP ($HP-trend$) has a positive and signicant effect on the probability of the ICC holding. All the other variables obtain insignificant coefficients. The trend in GDP has a positive and significant effect also on the probability of forming a cartel (κ_{it}), with a much larger value in the ICC equation. In addition, positive shocks to GDP affect cartel formation positively. Blue collar hours/gross value added, our (inverse) measure of labor productivity (see Table 4), has a negative impact on κ_{it}, suggesting that industries with a higher labor productivity are more likely to form a cartel. Our measure of variable costs of production, material share, has a negative effect on the probability of forming a cartel.

While the estimated observation probabilities (i.e., β^c_{it} and β^n_{it}, Table 5) are of no contemporary policy interest, let us note a few things about them. The probability of observing a cartel is decreasing both in the flow and in the stock of registered cartels, suggesting no (positive) learning. In contrast, the probability of observing a non-cartelized industry in a given year is increasing in the cumulative number of registered exits and decreasing in the number of exits registered in year $t-2$. Note also that because of the negative (positive) coefficient for the stock variables, the probability of observing that an industry does (not) have a cartel is decreasing (increasing) over time. This is in line with the empirical frequency displayed in Figure 3, suggesting that these parameters succeed in capturing the behavior of the Registry.

The results are rather similar when we fix the initial probability and when we estimate τ^n. However, we do reject the null hypothesis of $\tau^n = 0.5$. The estimated initial probability of not being in a cartel is 94%. This high probability may be explained by the fact that in 1951, the very strict war-time regulations that had been in place more or less since end of 1939 had only recently been lifted.

To further probe the robustness of our results, we have done the following.

First, our model necessitates the use of data on all industries, also those where the observation process always yields u. To compare our results to most of the existing work that only uses data on industries that have had an exposed cartel, we have re-estimated our HMM using data only on the industries with a known (= registered) cartel during the sample period. The results, omitted for brevity but available upon request, are very much in line with those of Table 3 - 5. This suggests that in our case, not using data on industries which do not have a known cartel would not bias the results greatly. Whether this extends to other data sets is naturally an open question.

Second, we checked whether our results are sensitive to $HP-trend$ entering linearly our specifications. A problem with using higher order terms is that they are very highly correlated (>0.9) with $HP-trend$. We therefore experimented by adding the square of the actual GDP series. We found no material changes to the reported results.

3.4 Policy Implications

Given that our data is from an era of legal cartels, we cannot identify the policy parameters σ_{it} and ν_{it}, or their effects on cartel stability. We can however identify the probability of forming a cartel (κ_{it}) and the probability that the ICC (H_{it}) holds. These are reported in Table 6 for the various models. We find that on average, κ_{it} is round 0.2. Recalling that we always reject Model 1 against Model 2 against Model 3, it seems that not using covariates biases this estimate downwards. The interpretation of this estimate is that on average, an industry that was not in a cartel last year has a 20% chance of being able to form a cartel this year.

In contrast, the estimated probability of the ICC holding (H_{it}) is on average very high, more than 0.9. The implication of this is that when cartels are legal, i) industries form a cartel with a very high probability *if* they get the chance and ii) that cartels, once formed, are very durable.

[Table 6 - H and κ]

[Figure 6 - Development of H and κ]

These predicted probabilities are allowed to vary over time in Models 2 and 3. In Figure 6 we show the development of predicted H_{it} and κ_{it} for Model 3 (using data from all industries, and estimated τ^n). The predicted probability of continuation is high and very stable with a weak positive trend. This trend is due to the trend in GDP. The opportunity probability (κ_{it}) varies more and also exhibits a positive trend which is stronger than that for H_{it}. The large increases in 1972-75 and 1989 are due to the large positive shocks in the aggregate demand in these periods (see $GDP - pos$ in Figure 5). Notice that κ_{it} is increasing trend-like, so even ignoring the effect of the positive GDP shocks, its value is significantly higher at the end of our sample period than at the beginning of it.[23]

These results suggest that the degree of cartelization may have increased over our sample period. We use the HMM structure of our model to illustrate this in two ways. First, using the local decoding method (see section 2.2.4 and Appendix C), we can analyze the hidden states and estimate the proportion of manufacturing industries that had a cartel in a given year. Second, we utilize the parameters of the hidden part of our model to perform a counterfactual analysis of the importance of the positive GDP shocks to the degree of cartelization.

The results of the local decoding exercise are displayed in Figure 7. The proportion of cartelized industries starts reasonably low at round 20%, reflecting the low value of κ_{it} in the early years. It then starts to increase, and jumps upwards in the early 1970s when κ_{it} increases both because it trends upwards, and because of the large positive GDP shock. We have one possible explanation for why the jump takes place in the early 1970s. Interestingly, the jump does coincide almost perfectly with the first oil crises, which also hit the open Finnish economy and many of its export oriented sectors. However, the export shock was positive as barter trade with the Soviet Union was important to many industries in Finland. Finland imported large amounts of oil, and exported mainly manufacturing goods to the Soviet Union. The trade between the Soviet Union and Finland was centrally negotiated and this forced industry managers to interact more frequently. This behavior would be consistent with the increase in κ_{it}.

[23] The development of H_{it} and κ_{it} is very similar when we use data on cartelized industries only. We overestimate κ_{it} somewhat when using data only from cartelized industries: The mean of κ_{it} is 0.254 using data on cartelized industries only, model 3 and estimated initial probabilities. This is 14% higher than the reported mean.

[Figure 7 - Estimated proportion of cartelized industries]

While the positive GDP shocks do play an important role, the upward trend in κ_{it} is much more important. This is so because of the high continuation probability H_{it}. It means that there was very little outflow from the stock of cartels. Cartels in almost all those industries that didn't previously have a cartel but formed one, did not collapse.[24] The relative importance of the trend in κ_{it} and the positive GDP shocks is illustrated in Figure 8, where we display the expected proportion of industries that have a cartel using both actual and counterfactual data. The proportion of cartelized industries is calculated by utilizing the fact that the hidden part of our model (i.e., the initial probability τ^n and the parameters of the transition matrix $\mathbf{A}_{it}$) allows us to recursively calculate the probability of a cartel existing in a given industry in a given period. We then average over industries.

We perform two counterfactuals. First, we replace the $GDP-pos$ values of the years 1972-75 with the average of all other years that had a positive shock. Second, we additionally replace also the large shock of 1989 with the average. The shocks of the early 1970s have a large impact in the short run, increasing the proportion of cartelized industries from 44% in 1971 to 92% in 1975. This rapid increase is followed by a small decrease. The effect of the very large 1989 GDP shock is small because the share of cartelized industries is already high by then. Noticeable is however that by end of the 1980s both counterfactual cases produce the same degree of cartellization as the actual data.

[Figure 8 - Counterfactual]

According to our estimates, over 90% of manufacturing industries were cartelized by the end of the sample. Inferring this directly from entries to

[24] To give an example, let's use 1970 numbers. In that year, 42% of industries had a cartel according to our estimates, meaning 98 of our 234 industries. Out of the 136 non-cartelized industries, 18% formed a cartel, and 93% of these satisfied their ICC, adding 23 cartels ($= 136 \times 0.18 \times 0.93$) to the cartel stock. At the same time, only 7 ($= 98 \times (1 - 0.93)$) previously established cartels died. Thus, the cartel stock went up from 98 to 114. In 1971, 8 cartels died, but 20 were born, taking the stock to 126. Even with the rather low 1970 value of κ_{it}, the proportion of industries that have a cartel would have increased rather rapidly because of the high value of H_{it}.

and/or exits from the Registry would have been impossible. Even if the welfare losses from these Finnish cartels were lower than the typical estimates in the literature, our estimate of the prevalence of cartels suggests that at least in Finland, there are high returns to effective competition policy.

4 Conclusions

The objective of this paper has been twofold: First, to build a model that allows one to estimate the parameters describing the efficiency of competition policy. The model uses industry (market) level information on cartels that is generated through modern competition policy actions and is thus becoming widely available. Second, to take a variant of the model to data on Finnish legal nationwide manufacturing cartels from 1951 to 1990 and to provide 1) estimates of the effects of industry characteristics and macro variables on the birth and death process of cartels and 2) an estimate of how cartelized Finnish manufacturing was during that era.

We show how the kind of data typically available on cartels, generated through competition policy actions, yields a Hidden Markov Model (HMM) once it is matched with the theoretical cartel model of Harrington and Chang (2009) and Chang Harrington (2009). HMMs are a well-studied and widely used class of statistical models that so far have found little use in industrial organization. The estimation approach is relatively flexible, as it can also be merged with other dynamic models of cartel behavior and modified to fit varying institutional environments. We chose the Harrington and Chang model because it endogenizes cartel births and deaths and incorporates the most important competition policy tools, namely CA detection and leniency. We specify the HMM to match the modern competition policy environment, derive the components of its likelihood function and discuss how the model could be estimated structurally, or using a reduced form approach. The former has the benefit that it allows for a counterfactual analysis of different competition policy regimes.

We take a variant of our HMM to data on Finnish legal cartels from 1951 to 1990. While necessarily uninformative about the effects of modern competition policy, these data allow us to uncover the (industry-specific) parameters that govern the hidden process of cartel births and deaths. These estimates yield information on the prevalence of cartels that we could expect if cartels were legal. Thereby, we are able to say something about the necessity of modern competition policy that i) declares cartels illegal, ii) seeks to detect cartels and ii) levies fines on detected cartels.

We find that both the probability of forming a cartel, and the probability to continue an existing cartel are increasing functions of GDP. Positive shocks to GDP affect the probability of forming a cartel, but not the incentive compatibility condition for continuing a cartel. Labor productivity has a positive impact on the probability of forming a cartel and variable costs a negative effect.

When we convert our parameter estimates to probabilities, we find that the mean probability of getting the chance to form a cartel is round 20%, while the probability of the ICC condition holding is as high as 90% on average. The former increases strongly over time while the latter is rather stable during our sample period. We estimate the proportion of Finnish manufacturing industries that were cartelized in our sample period and find that by the end of the period, almost all industries had a cartel. This result suggests that there is a high return to effective competition policy.

References

[1] Altman, Rachel, 2007, Mixed Hidden Markov Models: An Extension of the Hidden Markov Model to the Longitudinal Data Setting, *Journal of the American Statistical Association*, Vol. 102, No. 477, pp. 201 - 210.

[2] Asker, John, 2009, A Study of the Internal Organisation of a Bidding Cartel, *American Economic Review*, forthcoming.

[4] Bhar, Ramaprasad and Hamori, Shigeyuki, 2004, *Hidden Markov Models: Applications to Financial Economics*, Kluwer Academic Publishers.

[3] Bradburd, Ralph M., and Mead A. Over Jr., 1982, Organizational Costs, 'Sticky Equilibria,' and Critical Levels of Concentration, *Review of Economics and Statistics*, 64(1): 50–58.

[4] Brenner, Steffen, 2009, An Empirical Study of the European Corporate Leniency Program, *International Journal of Industrial Organization*, Vol. 27, pp. 639-645.

[5] Bresnahan, Timothy, 1989, Empirical Studies of Industries with Market Power, ch. 17. in (eds.) Schmalensee, R., and Willig, R. D., 1989, *Handbook of Industrial Organization*, Vol. I, Elsevier Science Publishers B.V.

[6] Cappé, Oliver, Moulines, Eric and Rydén, Tobias, 2005, *Inference in Hidden Markov Models*. Springer Series in Statistics, Springer.

[7] Chang, Myong-Hun and Harrington, Joseph E., Jr., 2009, The Impact of a Corporate Leniency Program on Antitrust Enforcement and Cartelization, mimeo, Johns Hopkins University.

[8] Ellison, Glenn, 1994, Theories of Cartel Stability and the Joint Executive Committee, *RAND Journal of Economics,* Vol. 25, No. 1, pp. 37-57.

[9] Engle, Clive and Hamilton, James D., 1990, Long Swings in the Dollar: Are They in the Data and Do Markets Know it?, *American Economic Review*, Vol. 89, pp. 689-713.

[10] Fellman, Susanna, 2009, The "Finnish Model of Capitalism" and Transforming Competition Policies, 1920s - 1990s, mimeo, University of Helsinki.

[11] Frass, Arthur G. and Greer, Douglas F., 1977, Market Structure and Price Colllusion: An Empirical Analysis, *Journal of Industrial Economics*, Vol. 26, No.1, pp. 21-44.

[12] Genesove, David and Mullin, Wallace, 1998, Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914, *RAND Journal of Economics,* Vol. 29, No. 2, pp. 355-377.

[13] Genesove, David and Mullin, Wallace, 2001, Rules, Communication, and Collusion: Narrative Evidence from the Sugar Institute Case, *American Economic Review*, Vol. 91, N0. 3, pp. 379 - 398.

[14] Green, Edward J. and Porter, Robert H., 1984, Noncooperative Collusion under Imperfect Price Information, *Econometrica*, Vol. 52, pp. 87-100.

[15] Haltiwanger, John C. and Harrington, Joseph E., Jr., 1991, The Impact of Cyclical Demand Movements on Collusive Behavior, *RAND Journal of Economics*, Vol. 22, pp. 89-106.

[16] Hamilton, James D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, Vol. 57, pp. 357-384.

[17] Harrington, Joseph E., Jr. and Chang, Myong-Hun, 2009, Modelling the Birth and Death of Cartels with an Application to Evaluating Antitrust Policy, *Journal of the European Economic Association*, Vol. 7, No. 6, pp. 1400-1435.

[18] Hay, George and Kelley, Daniel, 1974, An Empirical Survey of Price Fixing Conspiracies, *Journal of Law and Economics*, Vol. 17, pp. 13 - 38.

[19] Hodrick, Robert J. and Prescott, Edward C., 1997, Postwar U.S. Business Cycles: An Empirical Investigation, *Journal of Money, Credit and Banking*, Vol. 29, pp. 1-16.

[20] Jacquemin, Alexis, Tsuruhiko Nambu, and Isabelle Dewez, 1981, A Dynamic Analysis of Export Cartels: The Japanese Case, Economic Journal, Vol. 91, pp. 685–96.

[21] Judd, Kenneth L., and Su, Che-Lin, 2008, Constrained Optimization Approaches to Estimation of Structural Models, mimeo, Hoover Institution and Northwestern University.

[22] Lee, Lung-Fei and Porter, Robert H., 1984, Switching Regression Models with Imperfect Sample Separation Information - with an Application on Cartel Stability, *Econometrica,* Vol. 52, No. 2, pp. 391-418.

[23] Levenstein, Margaret C. and Suslow, Valerie Y., 2006a, What Determines Cartel Success? *Journal of Economic Literature,* Vol. XLIV (March 2006), pp. 43–95.

[24] Levenstein, Margaret C. and Suslow, Valerie Y., 2006b, Determinants of International Cartel Duration and the Role of Cartel Organization, WP. No. 1052, Ross School of Business, University of Michigan.

[25] MacDonald, Iain L. and Zucchini, Walter, 1997, *Hidden Markov Models and Other Discrete-valued TimeSeries*. Monographs on Statistics and Applied Probability 70, Chapman & Hall.

[26] MacDonald, Iain L. and Zucchini, Walter, 2009, *Hidden Markov Models for Time Series: An Introduction Using R*. Monographs on Statistics and Applied Probability 110, CRC Press / Taylor & Francis Group.

[27] Miller, Nathan, 2009, Strategic Leniency and Cartel Enforcement, *American Economic Review*, Vol. 99, No. 3, pp. 750–768.

[28] Pesendorfer, Martin, 2000, A Study of Collusion in First-Price Auctions, *Review of Economic Studies*, Vol. 67, pp. 381-411.

[29] Porter, Robert H., 1983, A Study of Cartel Stability: The Joint Executive Committee, 1880-1886, Bell Journal of Economics, Vol. 14, No. 2, pp. 301-314.

[30] Porter, Robert H., and Zona, J. Douglas, 1993, Detection of Bid Rigging in Procurement Auctions, *Journal of Political Economy*, Vol. 101, No. 3, pp. 518-538.

[31] Porter, Robert H., and Zona, J. Douglas, 1999, Ohio School Markets: An Analysis of Bidding, *RAND Journal of Economics*, Vol. 30, No. 2, pp. 263-288.

[32] Purasjoki, Martti and Jokinen, Juhani, 2001, *Kilpailupolitiikan Odotukset, Saavutukset ja Haasteet (Expectations, Achievements, and Challenges of Competition Policy)*, in Finnish only, Borenius & Kemppinen 90-vuotisjuhlajulkaisu.

[33] Rotemberg, Julio J. and Saloner, Garth, 1986, A Supergame-Theoretic Model of Price Wars During Booms, *American Economic Review,* Vol. 76, pp. 390-407.

[34] Röller, Lars-Hendrik and Steen, Frode, 2006, On the Workings of a Cartel: Evidence from the Norwegian Cement Industries, *The American Economic Review*, Vol. 96, pp. 321-338.

[35] Slade, Margaret E., 1989, Price Wars in Price-Setting Supergames, *Economica*, Vol. 56, pp. 295–310.

[36] Slade, Margaret E., 1990, Strategic Pricing Models and Interpretation of Price-War Data, *European Economic Review*, Vol. 34, No. 2-3, pp. 524–37.

[37] Suslow, Valerie Y., 2005, Cartel Contract Duration:Empirical Evidence from Inter-War International Cartels, *Industrial and Corporate Change*, Vol. 14, no. 5, pp. 705–44.

[38] Symeonidis, George, 2002, *The Effects of Competition*, MIT Press.

[39] Turner, Rolf, 2008, Direct Maximization of the Likelihood of a Hidden Markov Model, *Computational Statistics and Data Analysis*, Vol. 52, pp. 4147-4160.

[40] Wang, Peiming, and Alba, Joseph D., 2006, A Zero-Inflated Negative Binomial Regression Model with Hidden Markov Chain, *Economics Letters*, Vol. 92, pp. 209-213.

[41] Zucchini, Walter, Raubenheimer, David and MacDonald, Iain L., 2008, Modeling Time Seies of Animal Behaviour by Means of a Latent-State Model with Feedback, *Biometrics*, Vol. 64, pp. 807-815.

Appendix A: Finite HMM

To provide a formal definition for a HMM, let us assume that observations are recorded at equally spaced integer times $t = 1, 2, ..., T_i$ for cross-sectional units $i = 1, ..., N$. The observed data for i follow a HMM if the hidden states, $\{Z_{it}\}_{t=1}^{T_i}$, follow a Markov chain and if given Z_{it}, observation O_{it} at time t for unit i is independent of $O_{1t}, ..., O_{i,t-1}, O_{i,t+1}, ..., O_{iT_i}$ and $Z_{1t}, ..., Z_{i,t-1}, Z_{i,t+1}, ..., Z_{iT_i}$. This property means that in a standard HMM, the observations are independent conditional on the sequence of hidden states.

The general econometric/statistical theory and scope of applications of the HMMs is broad (see, e.g., Cappé, Moulines and Rydén 2005, Zucchini and MacDonald 2009 and, for a seminal reference, MacDonald and Zucchini 1997, on which this section builds), but for the purposes of our analysis, we can focus on the case in which Z_{it} takes on values from a finite set (state space), $S_Z = \{s_1, s_2, ..., s_{\bar{Z}}\}$, where $\bar{Z}$ is known. We also assume that Y_{it} is a discrete (categorical) random variable, taking on values from a finite (observations) set, $S_O = \{o_1, o_2, ..., o_{\bar{O}}\}$, where $\bar{O}$ is known. We define $\mathbf{O}_i$ to be the T_i -dimensional vector of observations on i and $\mathbf{O}$ the $\sum_{i=1}^{N} T_i$ -dimensional vector of all observations. The vectors of hidden states, $\mathbf{Z}_i$ and $\mathbf{Z}$, are defined similarly. Finally, we let $\mathbf{x}_{it}$ denote the K-dimensional vector of covariate values of unit i at t, with $\mathbf{x}_i = \{\mathbf{x}_{i1}, ..., \mathbf{x}_{iT_i}\}$.

The HMM is fully specified by the initial and transition probabilities of the hidden Markov chain and by the distribution of O_{it}, given Z_{it}. For a cross-sectional unit i, these three stochastic elements can be specified as follows:

First, we specify the probability that unit i is at the unobserved state $k \in S_Z$ in the initial period (i.e., $Z_{i1} = k$), given its contemporary covariate values. These initial state probabilities are denoted

$$\tau_i^k = P\left(Z_{i1} = k \left| \mathbf{x}_{i1} \right.\right). \tag{12}$$

Second, we specify the (hidden) transition probabilities. They give the probability that unit i is at state $k \in S_Z$ in period t, given that it was at state $j \in S_Z$ in period $t-1$, and given its contemporary covariate values. These transition probabilities are

$$a_{it}^{jk} = P(Z_{it} = k | Z_{i,t-1} = j, \mathbf{x}_{it}). \tag{13}$$

This formulation shows that we allow the Markov chain to be non-homogenous (i.e., the transition probabilities can depend on a time index) and that conditional on $\mathbf{x}_{it}$, the current state depends only on the previous state (the Markov property).[25]

The third stochastic element of the HMM are the observation probabilities ('state-dependent probabilities'). The observation probabilities give the probability of observing $w \in S_O$ when the unobserved state is $k \in S_Z$ at t, i.e.,

$$b_{it}^{k}(w) = P(O_{it} = w | Z_{it} = k, \mathbf{x}_{it}). \tag{14}$$

This formulation shows that $b_{it}^{k}(w)$ can depend on covariates and that conditional on $\mathbf{x}_{it}$, the observation at time t depends only on the current hidden state and is independent of the previous observations (and states).

To derive the likelihood of the HMM, let Θ denote the model parameters, $\mathbf{D}_{i1}$ the $(\bar{O} \times 1)$ vector with elements $d_{i1}^{k}(w) = \tau_i^k b_{i1}^{k}(w)$, $\mathbf{D}_{it}$the $(\bar{O} \times \bar{O})$ matrix with elements $d_{it}^{jk}(w) = a_{it}^{jk} b_{it}^{k}(w)$ for $t > 1$, and $\mathbf{1}$ the $(\bar{O} \times 1)$ vector of ones. As shown in e.g. MacDonald and Zucchini (2009, p. 37), Altman (2007) and Zucchini, Raubenheimer and MacDonald (2008), the likelihood for the whole observed data can be written as

$$L(\Theta; \mathbf{o}) = \prod_{i=1}^{N} \left\{ (\mathbf{D}_{i1})' \left(\prod_{t=2}^{T_i} \mathbf{D}_{it} \right) \mathbf{1} \right\} \tag{15}$$

where $\mathbf{o}$ denotes the data (the realization of $\mathbf{O}$).

Appendix B: Hypothetical Cartel Data

Table B.1 illustrates the type of *observed* data a cartel researcher might have access to. For this hypothetical example, we set $T = 5$ and use the following notation for the observed states: "Not in a cartel" $= n$, "In a cartel" $= c$,

[25] If the hidden process is stationary and has a unique stationary distribution, the stationary probabilities can be used as the initial probabilities. The stationary distribution for the Markov chain does not necessarily exist if one or more of the covariates are functions of time. If there is no stationary distribution, an initial distribution for Z_{i1} can be assumed and the parameters of this distribution can be treated as nuisance parameters. It may, for example, be sensible to assume that the initial probabilities do not depend on the covariates and are the same for all for cross-sectional units.

"Detected and shut down by the CA" = d, "Leniency" = l and "Unknown / unobserved" = u.

Table B.1: Hypothetical cartel data

time/industry	1	2	3	4	5	6	...	N
$t = 1$	u	u	c	c	u	u	...	u
$t = 2$	u	n	c	c	n	u	...	u
$t = 3$	u	c	d	c	n	u	...	u
$t = 4$	d	d	u	l	u	u	...	u
$t = 5$	u	u	u	u	u	u	...	u

The (hypothetical) data tell us (see column 1), for example, that for industry 1, $\mathbf{y}_1 = (u, u, u, d, u)'$. This industry had a cartel in period $t = 4$ that was detected and shut down by the CA during that period. The records provide no reliable information about its status prior to or after the detection. Industry 2 had a cartel in period $t = 4$ that was detected and shut down by the authorities during that period. For this industry, the cartel investigations by the authorities (and thereby their records) reveal that the cartel had been up and running for one year prior to its detection, and the court established that no cartel existed two years before the detection took place. However, the records provide no realiable information about the status of the industry for period $t = 1$ or the post-detection period $t = 5$. Industry 3 had a cartel in period $t = 3$ that was detected and shut down by the authorities during that period. For this industry, the cartel investigations by the authorities reveal that the cartel had been up and running for two years prior to its detection. But the sample starts at $t = 1$, so this industry enters the data in a cartel. The records allow one to determine these facts, but provide no realiable information about the status of the industry for the post-detection periods $t = 4$ and $t = 5$.

For industry 4, the data are informative about one usage of the leniency facility ($t = 4$). The investigations then revealed that the industry was in a cartel for three years prior to a member applying for leniency. Industry 5 is an example of what may be rarely observed in cartel data: This case corresponds to an industry that was suspected and investigated (and perhaps taken into a court) for having a cartel over a two-year period. The records (e.g., the court decision) show, however, that it eventually turned out that the industry had no cartel.

For the remaining industries (i.e. for $i = 6, ..., N$ in our hypothetical example), the (published) records of the CA or courts provide no realiable information about their status, perhaps because they have never been investigated for having a cartel or perhaps because they were suspected of having one, but the evidence was too weak to result in a published cartel case. These additional industries can, however, be included in the data set because we assume (as does Miller 2009) that the number of the industries in the given sector of the economy from which the detected cartel cases come can be identified. If nothing else is available, the number of industries covered by the industrial statistics of the statistical office can be used to identify N.

Appendix C: State Prediction and Learning

In the established literature on HMMs there are three primary methods that can be used to analyze the hidden states. They are state prediction, and local and global decoding. To give a formal description of them, we build on Zucchini and MacDonald (2009) and introduce some new notation. To this end, let $\mathbf{O}_i^{(t)} \equiv (O_{i1}, O_{i2}, ..., O_{it})$ denote observation history of industry i from time 1 to t, with corresponding realization $\mathbf{o}_i^{(t)}$. Similarly, let $\mathbf{O}_i^{(t+1,T)} \equiv (O_{i,t+1}, ..., O_{iT_i})$ denote 'future' from $t+1$ to T_i , with corresponding realization $\mathbf{o}_i^{(t+1,T)}$. We further define $L_{T_i} = (\mathbf{D}_{i1})^{'} \left(\prod_{t=2}^{T_i} \mathbf{D}_{it}\right) \mathbf{1}$. Finally, we need two $(1 \times \overline{Z})$ vectors, called the *forvard* and *backward* probability vectors. For $t = 1, ..., T_i$, the former is defined by

$$\zeta_{\mathbf{i},\mathbf{t}} \equiv (\mathbf{D}_{i1})^{'} \left(\prod_{s=2}^{t} \mathbf{D}_{is}\right).$$

This vector has property $\zeta_{\mathbf{i},\mathbf{t}} = \zeta_{\mathbf{i},\mathbf{t-1}} \mathbf{D}_{it}$ and its k^{th} element, $\zeta_{i,t}(k)$, is the joint probability $Pr\left[\mathbf{O}_i^{(t)} = \mathbf{o}_i^{(t)}, Z_{it} = k\right]$.

The vector of backward probabilities is in turn defined by

$$\epsilon_{\mathbf{i},\mathbf{t}}{}^{\prime} \equiv \left(\prod_{s=t+1}^{T_i} \mathbf{D}_{is} \mathbf{1}\right).$$

The j^{th} component of $\epsilon_{\mathbf{i},\mathbf{t}}{}^{\prime}$ is denoted $\epsilon_{i,t}(k)$ and it is equal to the conditional probability $Pr\left[\mathbf{O}_i^{(t,T)} = \mathbf{o}_i^{(t,T)} | Z_{it} = k\right]$. It can be shown that $L_{T_i} = \zeta_{\mathbf{i},\mathbf{t}} \epsilon_{\mathbf{i},\mathbf{t}'} = \zeta_{i,T_i} \mathbf{1}$.

With this notation at our disposal, we can determine the first quantity of interest to us. It is state prediction (i.e., the conditional distribution of state Z_{it}) for some $t > T_i$ and i. Provided that one has a view on which values the covariates take on during the future time period of interest (i.e., $\mathbf{x}_{i,T+1}$), this conditional probability can be computed as

$$Pr\left[Z_{i,T+1} = c \,\middle|\, \mathbf{O}_i^{(T)} = \mathbf{o}_i^{(T)},\ \mathbf{x}_{i,T+1}\right] = \frac{\zeta_{i,T}\mathbf{a}_{i,T_i+1}(c)}{L_{T_i}},$$

where we explicitly allow for the depence of the state prediction on the predicted values of the covariates and where vector $\mathbf{a}_{i,T_i+1}(c)$ is the column of $\mathbf{A}_{it}$, evaluated at $\mathbf{x}_{i,T+1}$. The elements of this vector correspond to transitions from $Z_{i,t-1} = n$, $Z_{i,t-1} = c$, or $Z_{i,t-1} = d$ to $Z_{it} = c$. A similar expression can be derived for the probability that the industry is not in a cartel at $T_i + 1$.

The second quantity of interest is based on local decoding. For any $t \in \{1, ..., T_i\}$, the interest is in finding the state that is most likely to have generated the observed data. For t, this most probable state, k_t^*, is

$$k_t^* = argmax_{k=1,...,\bar{Z}} Pr\left[Z_{it} = k \,\middle|\, \mathbf{O}_i^{(T)} = \mathbf{o}_i^{(T)}\right]$$

where $Pr\left[Z_{it} = k \,\middle|\, \mathbf{O}_i^{(T)} = \mathbf{o}_i^{(T)}\right] = \zeta_{i,t}(k)\epsilon_{i,t}(k)/L_{T_i}$. This expression allows one to determine, for example, whether it is more likely that industry i was in a cartel in period t than that it was not for those parts of the observed data that are uninformative about the status of this particular industry.

The third method is global decoding. It looks for the entire sequence of states, $\mathbf{z}_i^{(T)}$, which maximizes

$$Pr\left[\mathbf{Z}_i^{(T)} = \mathbf{z}_i^{(T)} \,\middle|\, \mathbf{O}_i^{(T)} = \mathbf{o}_i^{(T)}\right]$$

where $\mathbf{Z}_i^{(t)}$ and $\mathbf{z}_i^{(t)}$ denote state histories. There is a dynamic programming algorithm, called the Viterbi algorithm, which can be used to find the optimal sequence for industry i.

Tables and Figures

Table 1: Observed transitions

Industries with a registered cartel				
	n_t	c_t	u_t	Row total
n_{t-1}	207	65	89	361
	(57.34)	(18.01)	(24.65)	(100)
c_{t-1}	78	312	184	574
	(13.59)	(54.36)	(32.06)	(100)
u_{t-1}	80	186	3050	3316
	(2.41)	(5.61)	(91.98)	(100)
Column total	365	563	3323	4251
	(8.59)	(13.24)	(78.17)	(100)
All industries				
	n_t	c_t	u_t	Row total
n_{t-1}	207	65	89	361
	(57.3)	(18.0)	(24.7)	(100)
c_{t-1}	78	312	184	574
	(13.6)	(54.4)	(32.1)	(100)
u_{t-1}	80	186	7925	8191
	(1.0)	(2.3)	(96.8)	(100)
Column total	365	563	8198	9126
	(4.0)	(6.2)	(89.8)	(100)

NOTES: reported numbers are # observations and (%) of observations on the row.

Table 2: Descriptive statistics

Variable	#Obs	Mean	S.D.	Min.	Max.
$HP-trend$	9360	722.13	308.73	297.34	1317.29
$GDP-neg$	9360	6.30	9.91	0	38.89
$GDP-pos$	9360	6.24	11.23	0	42.43
GVP	9216	2809560	3002387	4413.7	1.26E+07
$Plants$	9216	452.52	376.06	6	1602
$Hours/GVP$	9216	0.021	0.017	0	0.172
$Materialshare$	9216	0.573	0.137	0.122	0.919
$Death-stock$	9360	158.75	172.33	0	581
$Death-flow$	9360	22.5	16.378	0	72
$Birth-stock$	9360	423.23	320.76	0	900
$Birth-flow$	9360	14.53	14.02	0	47

NOTES: The number of observations is lower for industry variables than others due to missing observations.

Table 3: Estimation results H

H	$\tau^n = 0.5$ M1	$\tau^n = 0.5$ M2	$\tau^n = 0.5$ M3	Estimated τ^n M1	Estimated τ^n M2	Estimated τ^n M3
$HP-trend$	-	0.0010	0.0010***	-	0.0009***	0.0009***
		(0.0002)	(0.0002)		(0.0002)	(0.0002)
$GDP-neg$	-	0.0013	0.001	-	0.0008	0.0010
		(0.0043)	(0.0043)		(0.0043)	(0.0043)
$GDP-pos$	-	0.0030	0.0041	-	0.0029	0.0039
		(0.0041)	(0.0043)		(0.0041)	(0.0043)
GVP	-	-	-9.83E-09	-	-	-9.42E-09
			(2.18E-08)			(2.16E-08)
$Plants$	-	-	-0.00004	-	-	-0.00005
			(0.00015)			(0.00014)
$Hours/GVP$	-	-	0.34208	-	-	0.16835
			(0.51121)			(4.22278)
$Materialshare$	-	-		-	-	0.37545
						(0.50781)
$Constant$	1.9919***	0.9359***	0.7880**	1.9976***	0.9995***	0.8247***
	(0.0410)	(0.1721)	(0.3616)	(0.0410)	(0.1706)	(0.3617)

NOTES: Reported numbers are coefficient and (s.e.). ***, **, and * denote significance at 1, 5, and 10% level.

Table 4: Estimation results κ

κ	$\tau^n = 0.5$ M1	$\tau^n = 0.5$ M2	$\tau^n = 0.5$ M3	Estimated τ^n M1	Estimated τ^n M2	Estimated τ^n M3
$HP-trend$	-	0.0027***	0.0024***	-	0.0027***	0.0024***
		(0.0002)	(0.0002)		(0.0002)	(0.0002)
$GDP-neg$	-	-0.0037	-0.0043	-	-0.0043	-0.0049
		(0.0055)	(0.0055)		(0.0054)	(0.0055)
$GDP-pos$	-	0.0380***	0.0371***	-	0.0378***	0.0372***
		(0.0045)	(0.0046)		(0.0045)	(0.0046)
GVP	-	-	2.49E-08	-	-	2.93E-08
			(2.8E-08)			(2.77E-08)
$Plants$	-	-	0.0001	-	-	0.0001
			(0.0001)			(0.0001)
$Hours/GVP$	-	-	-8.7053**	-	-	-8.8191**
			(3.7926)			(3.7169)
$Materialshare$	-	-	-1.3245***	-	-	-1.5006***
			(0.4404)			(0.4291)
$Constant$	-1.3252***	-3.2184***	-2.1980***	-1.3251***	-3.1878***	-2.0843***
	(0.0324)	(0.1146)	(0.3248)	(0.0322)	(0.1123)	(0.3152)

NOTES: Reported numbers are coefficient and (s.e.). ***, **, and * denote significance at 1, 5, and 10% level.

Table 5: Estimation results β's

β^c	$\tau^n = 0.5$			Estimated τ^n		
	M1	M2	M3	M1	M2	M3
Birth − stock	-0.0203***	-0.0158***	-0.0159***	-0.0210***	-0.0173***	-0.0175***
	(0.0021)	(0.0022)	(0.0023)	(0.0021)	(0.0022)	(0.0022)
Birth − flow	-0.0029***	-0.0036***	-0.0036***	-0.0030***	-0.0037***	-0.0037***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	0.6555***	1.0631***	1.0577***	0.7078***	1.1675***	1.1680***
	(0.0903)	(0.1036)	(0.1039)	(0.0891)	(0.1053)	(0.1052)
β^n						
Death − stock	0.0220***	0.0182***	0.0183***	0.0219***	0.0182***	0.0183***
	(0.0018)	(0.0011)	(0.0011)	(0.0018)	(0.0011)	(0.0011)
Death − flow	-0.0156**	-0.0146**	-0.0149**	-0.0156**	-0.0146**	-0.0149**
	(0.0080)	(0.0068)	(0.0068)	(0.0080)	(0.0068)	(0.0068)
Constant	-2.2456***	-2.3557***	-2.3520***	-2.2483***	-2.3575***	-2.3542***
	(0.0608)	(0.0616)	(0.0617)	(0.0608)	(0.0616)	(0.0616)
$\Phi^{-1}(\tau^n)$	0	0	0	1.6538***	1.6226***	1.6439***
				(0.1419)	(0.1421)	(0.1424)
LL.	-3052.5	-2773.8	-2752.6	-2940.7	-2666.2	-2643.5
Nobs.	9360	9360	9216	9360	9360	9216
LR (M2 vs. M1)	-	557.48 (6)	-	-	548.9 (6)	
LR (M3 vs. M2)	-	-	42.47 (8)	-	-	45.36 (8)
LR ($\tau^n = 0.5$ vs. est'ed. τ^n)	-	-	-	223.78 (1)	215.16 (1)	218.04 (1)

NOTES: Reported numbers are coefficient and (s.e.). ***, **, and * denote significance at 1, 5, and 10% level. ***, **, and * denote significance at 1, 5, and 10% level. LR = likelihood ratio test value, reported numbers test value and (d.f.).

Table 6: Policy parameters

Parameter	$\tau^n = 0.5$			Estimated τ^n		
	M1	M2	M3	M1	M2	M3
H	0.977	0.944	0.943	0.977	0.946	0.946
κ	0.093	0.228	0.224	0.093	0.226	0.222

NOTES: Reported numbers are means of the estimated values.

Figure 1 The timeline and our definition of cartel status

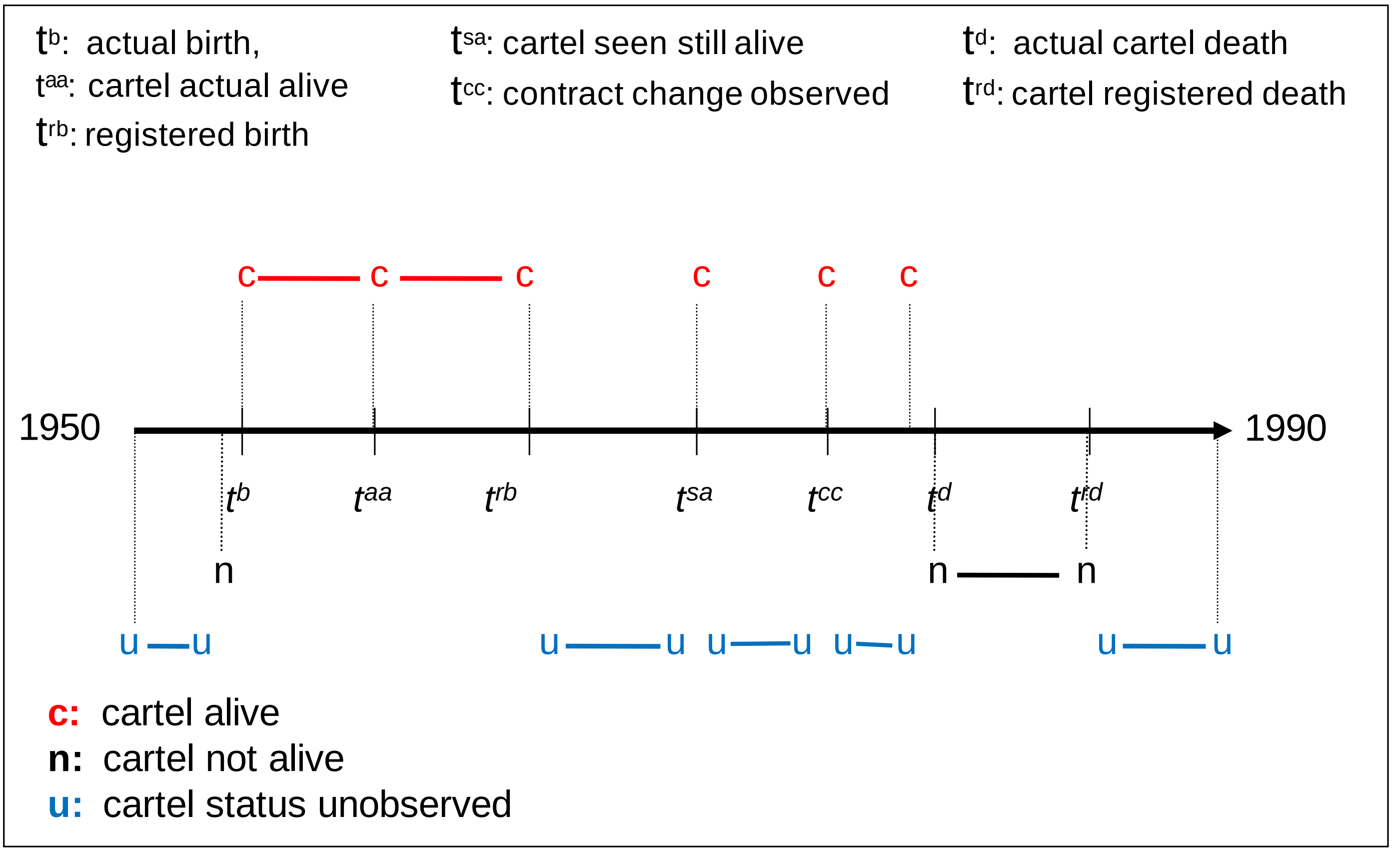

Figure 2 Kernel density of real and registry cartel survival time (kernel = epanechnikov, bandwidth = 3.7839)

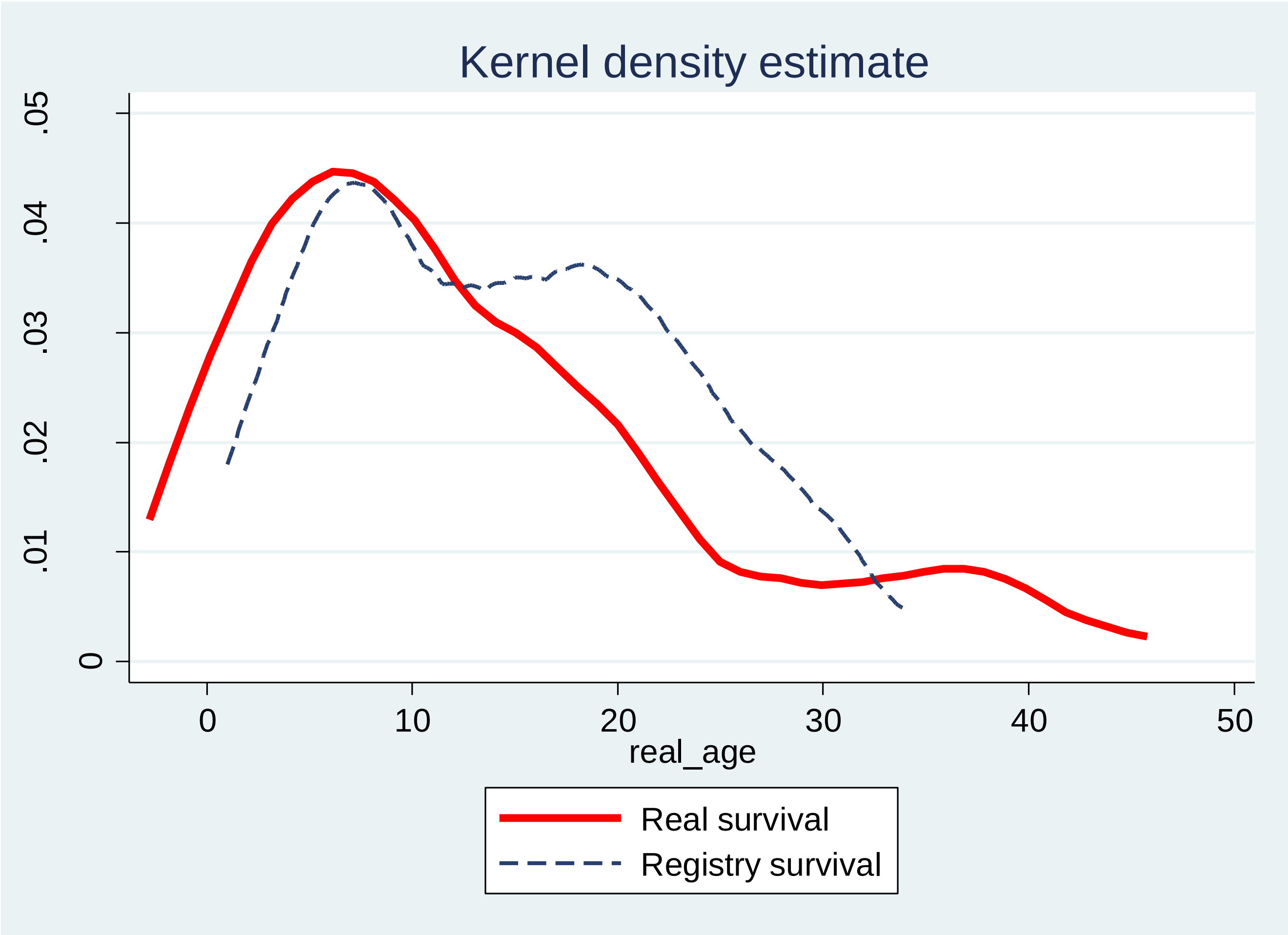

Figure 3 The time variation of states

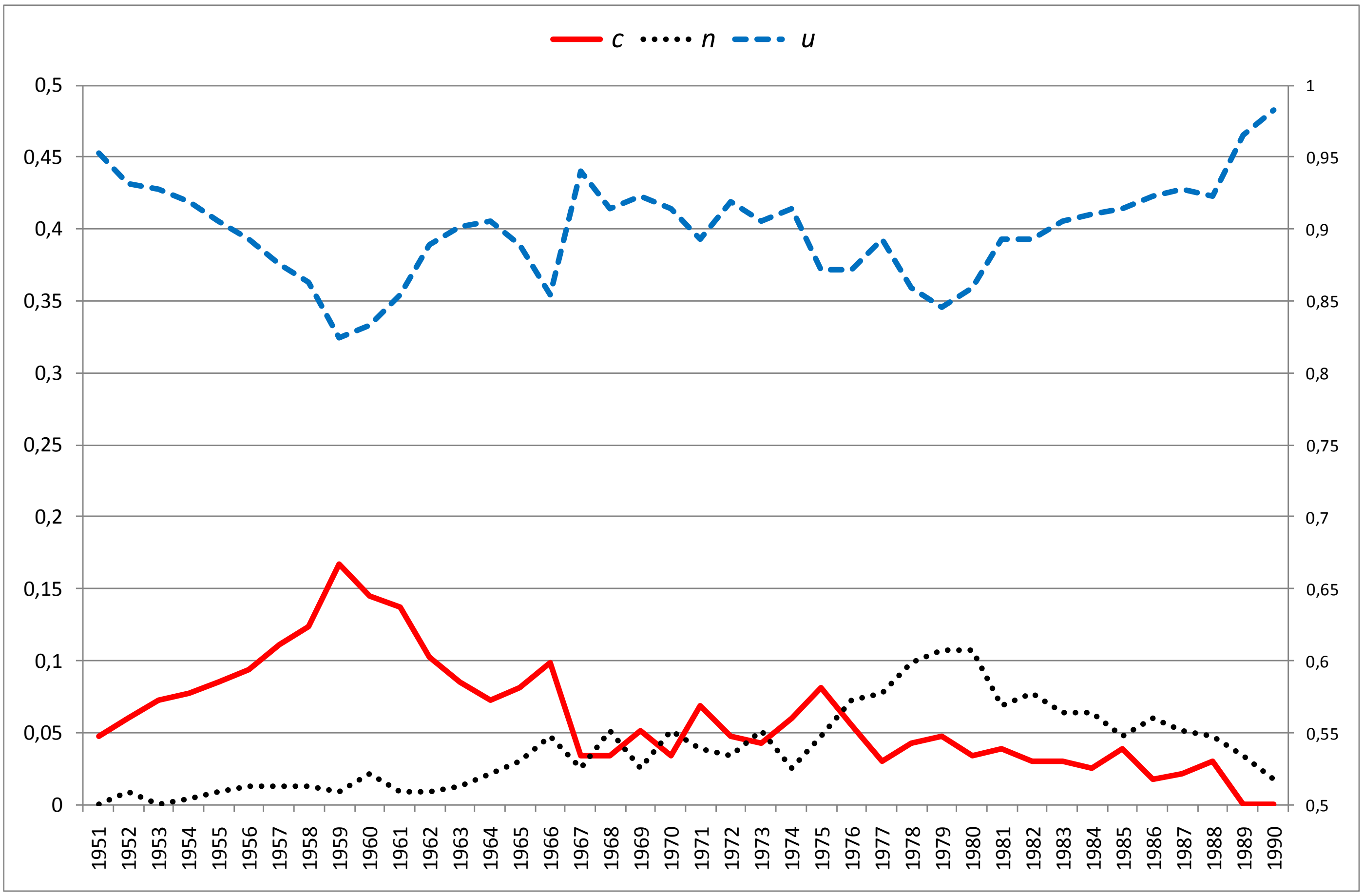

Figure 4 Registry deaths and births over time

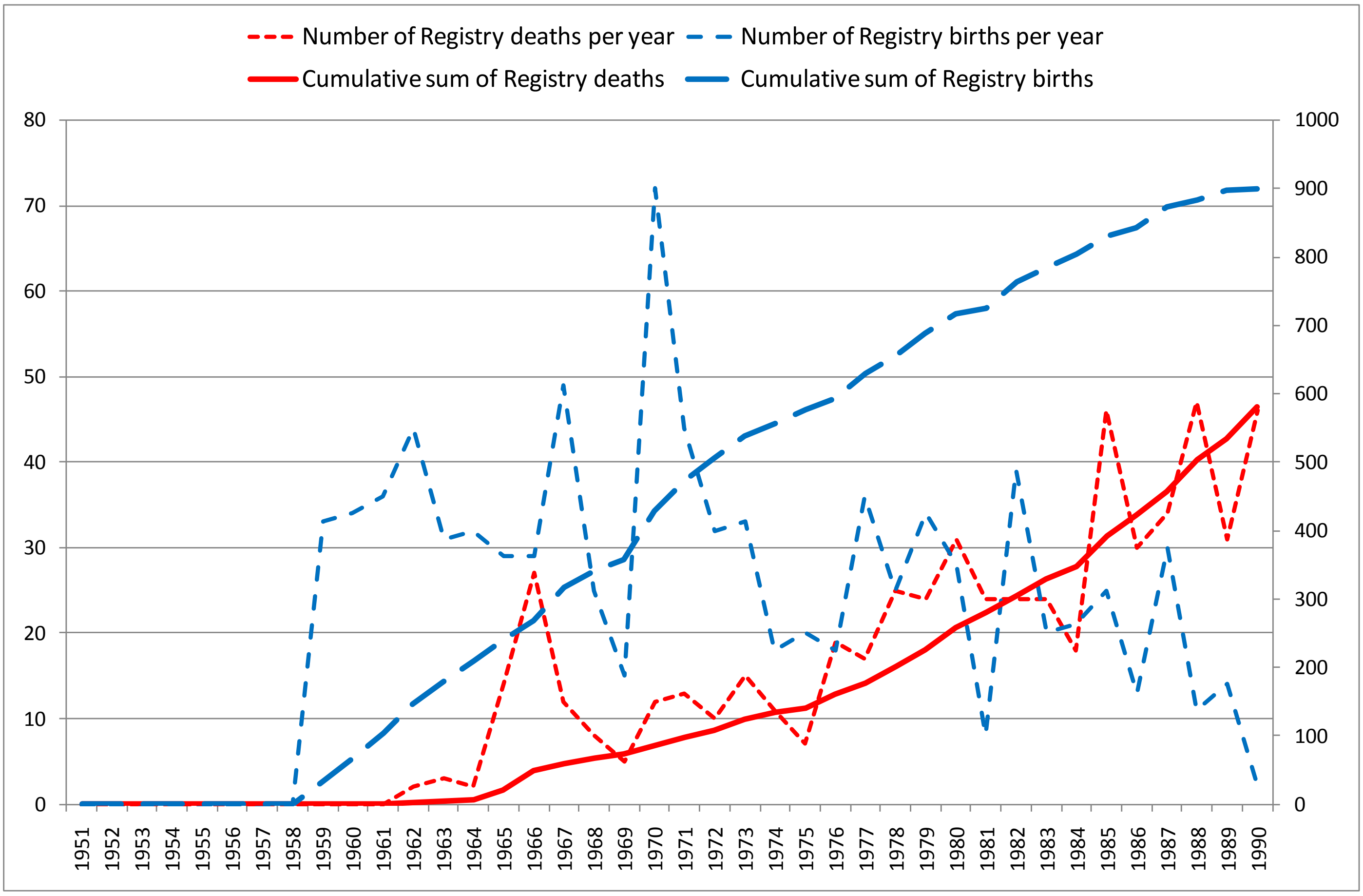

Figure 5 The development in Finnish GDP, long run HP-trend and deviations from trend

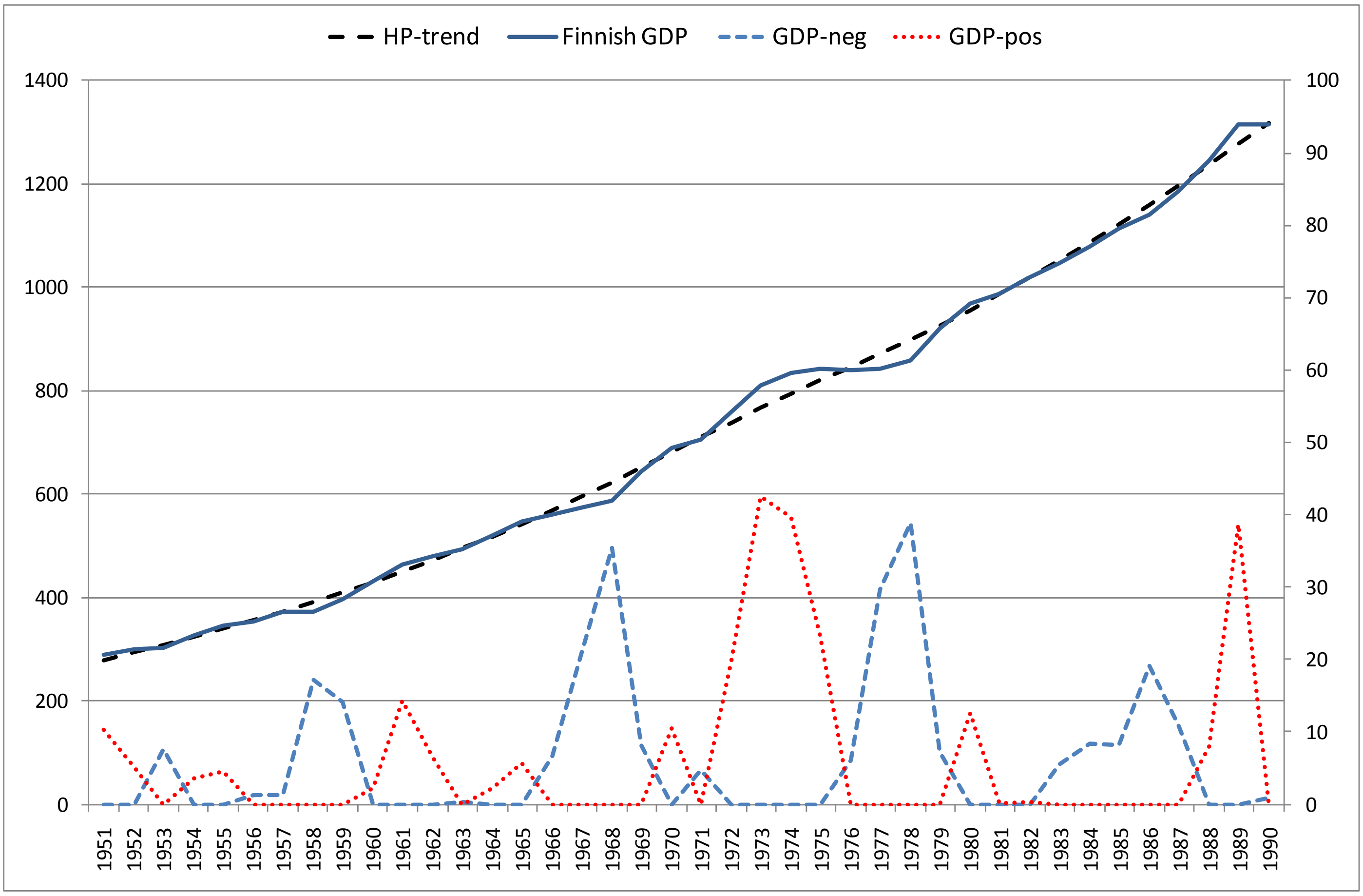

Figure 6 The time variation in H and κ

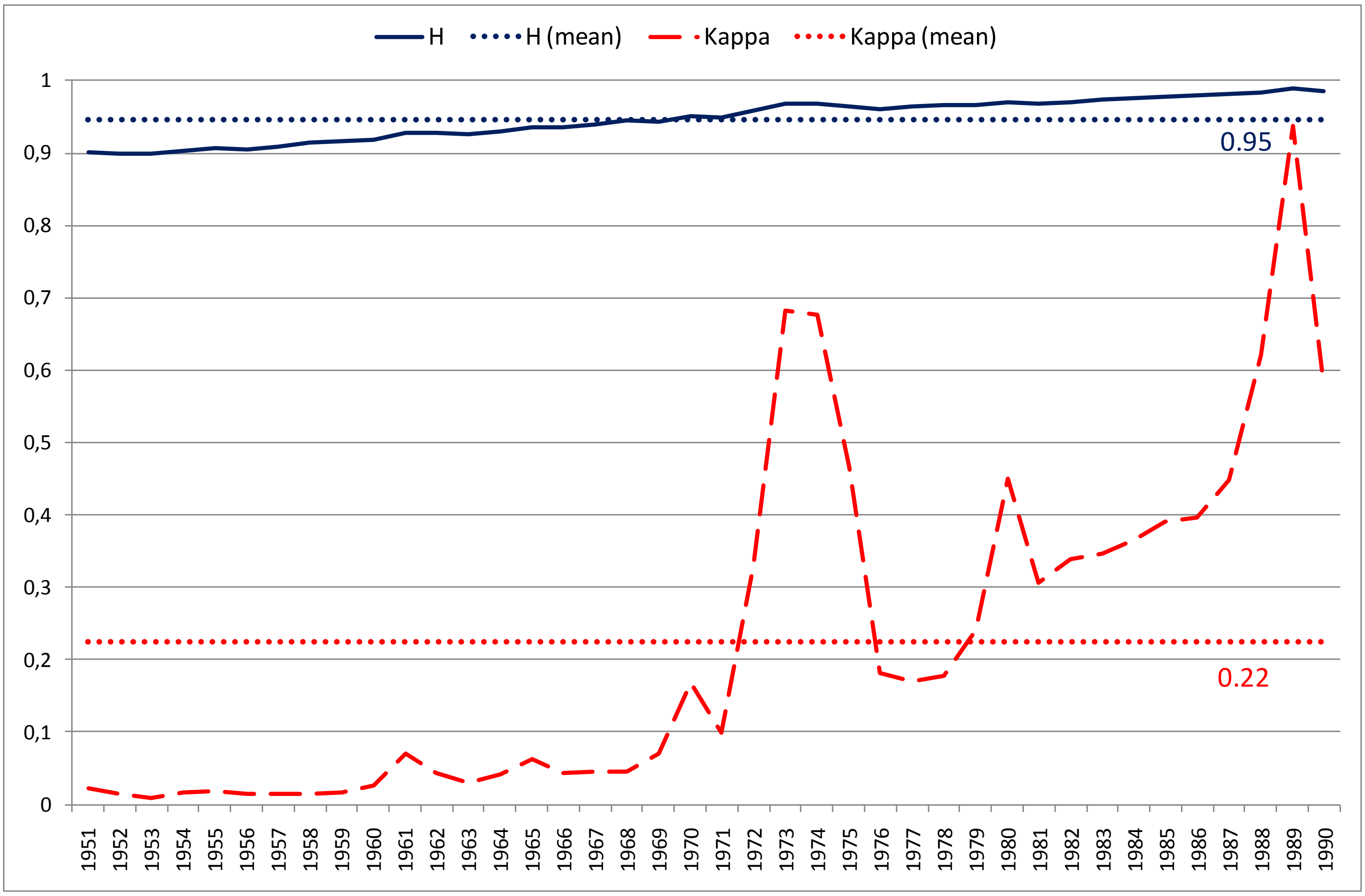

Figure 7 Local decoding of industry probabilities 1951-1990

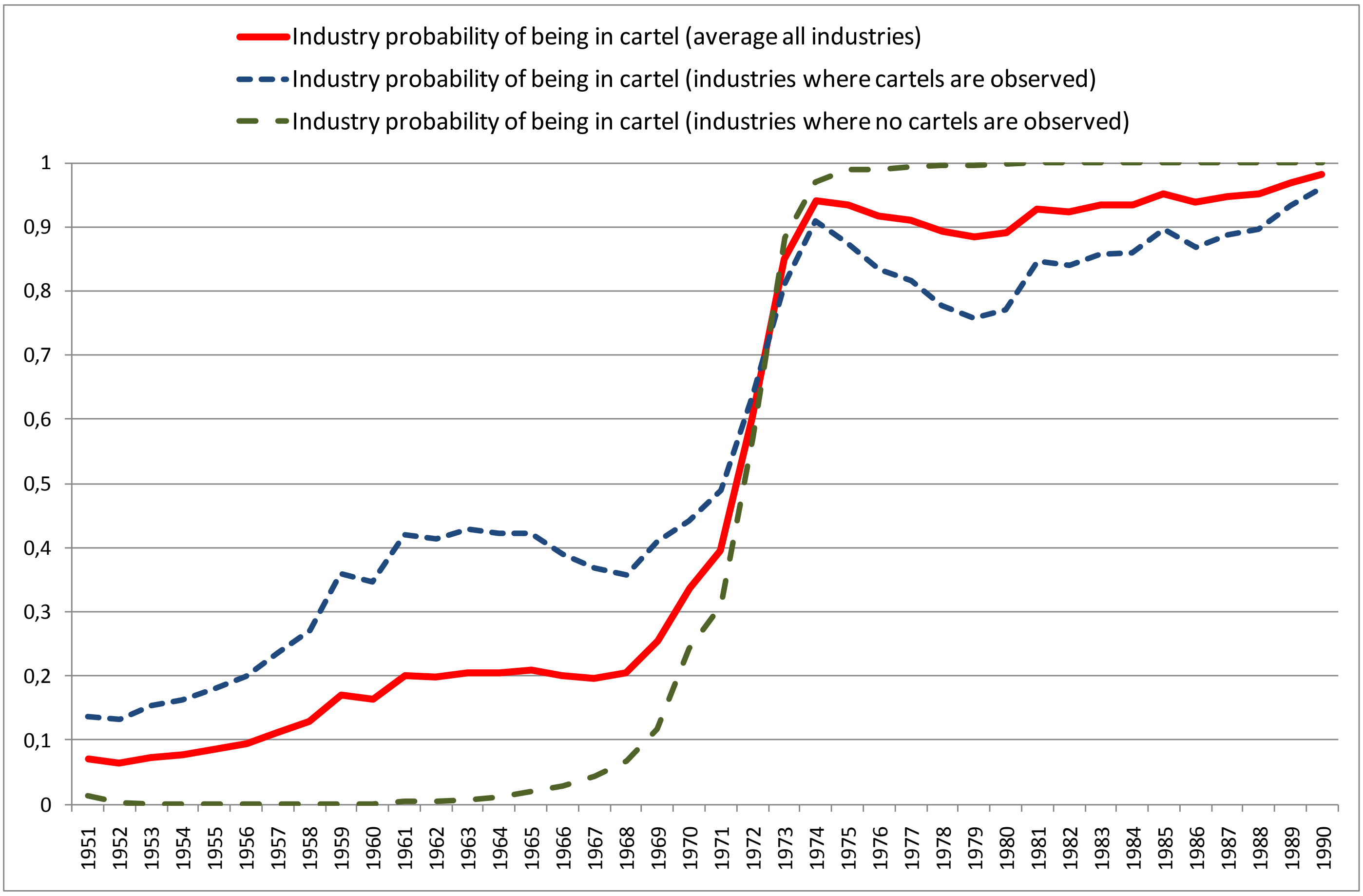

Figure 8 Counterfactual analysis implementing different GDP shock scenarios