Online Appendix to: Cartels Uncovered

Ari Hyytinen
University of Jyväskylä

Frode Steen
NHH Bergen & CEPR

Otto Toivanen
Aalto U. School of Business, KU Leuven & CEPR

January 11, 2018

Appendices

[Online appendix, not intended for publication]

Appendix A: Descriptive statistics

In this appendix, we report the count of c- and n -observations in the estimation data (Figure A1) and compare them the estimated observation probabilities, time series of Finnish GDP (decomposed to HP-trend and to positive and negative shocks) over 1951-1990 (Figure A2), and the descriptive statistics of the explanatory variables (Table A1).

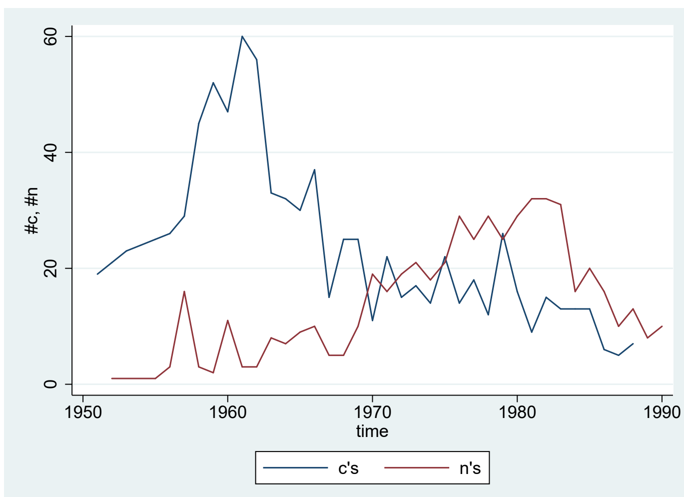

Figure A1: Count of c- and n- observations in the estimation data

Comparing Figure A1 to our estimated observation probabilities (Figure 2) we note the following: as explained in the main text, during the early years of the Registry, the legislation was on purpose lenient on cartels so as to facilitate registrations. Towards the end of our sample period some of the cartels probably started avoiding registration despite the more stringent legal requirements for doing so. It also is plausible that, to the extent they existed, the Registry was very keen to register at least all the major manufacturing cartels early on and that the Registry started paying more systematic attention to removing the ceased cartels from the Registry only later. Consistent with this, the law change in 1973 is likely to have increased the likelihood of observing n's, as it required more systematic reporting of cartel deaths. This suggests that even though there were some concerns about the ability of the Registry to get cartels registered during the early years of the Registry's existence, this was less of an issue for the nationwide manufacturing cartels. In the early 1960s, β^c starts to decline (see Figure 2), meaning that a lower proportion of observations in hidden state c are observed to be in that state. This means that the hidden and observed c-series start to diverge. A similar but reverse story holds for the n-states. These patterns of the observation process are consistent with the view that the nationwide manufacturing cartels had initially few reasons to hide their activity and that the athmosphere changed towards the end of our sample period, when the incentives of such cartels to disclose their ongoing activities diminished and to report their ceased activities increased.

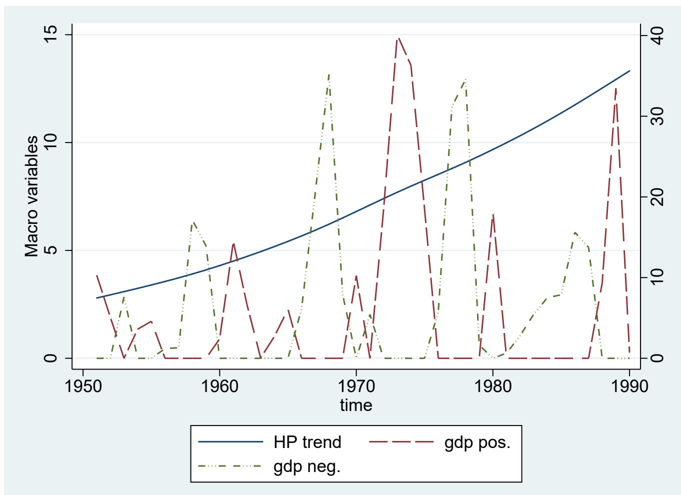

Figure A2: Macro shocks and GDP trend

Table A1: Descriptive statistics

Variable	Definition	Mean	Median	S.d.
$Hp - trend$	GDP volume index / 100, original series 100 = 1913	7.271	6.954	3.144
$Gdp - pos$	$\lvert \text{GDP vol. - } HP - trend \text{ in year } t \rvert \mid \text{GDP vol.} > HP - trend$	6.027	0	10.325
$Gdp - neg$	$\lvert \text{GDP vol. } HP - trend \text{ in year } t \rvert \mid \text{GDP vol.} < HP - trend$	6.082	0.920	9.556
$Law - index$	0 before 1959 law; increase by 1 at each law change	1.200	1	0.980
$Homog - d$	1 if homogenous goods, 0 otherwise	0.378	0	0.485
$Death - stock$	#exits from the Registry by $t - 1$	1.387	0.79	1.594
$Death - flow$	#exits from the Registy in $t - 1$	13.375	11.5	13.292
$Birth - stock$	#entries into the Registry by $t - 1$	4.007	3.93	3.181
$Birth - flow$	#entries into the Registy in $t - 1$	22.45	25	1.644
$Birth - count$	#entries into the Registry in industry i by $t - 1$	0.882	0	1.438

Notes: Estimation of the $HP - trend$ was done using a smoothing index of 100.

Appendix B: Illegal cartels and modern competition policy

In this appendix, we briefly describe how a HMM can be specifed so that it allows modeling discovered cartels during an era when cartels are illegal. We start with a brief description of a finite HMM, as by doing so we can define the notation used later in this appendix. We then illustrate what kind of observations on cartels modern data sets are likely to have and how the hidden process and the observation process can be specified to match the institutional environment.

Basic Description of a finite HMM

Let observations be 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 (see e.g. Cappé, Moulines and Rydén 2005, Zucchini and MacDonald 2009, on which this description builds). The hidden states, Z_{it}, take on values from a finite set (state space), $S_Z = (s_1, s_2, ..., s_{\bar{Z}})$, where $\bar{Z}$ is known and observations, O_{it}, are a discrete (categorical) random variable, taking on values from a finite (observation) set, $S_O = (o_1, o_2, ..., o_{\bar{O}})$, where $\bar{O}$ is known. Finally, 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}. These three stochastic elements are as follows: First, $\tau_i^k = P\left(Z_{i1} = k \left| \mathbf{x}_{i1} \right.\right)$ is the initial state 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. Second, the (hidden) transition probabilities 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 covariate values. These transition probabilities are denoted $a_{it}^{jk} = P(Z_{it} = k | Z_{i,t-1} = j, \mathbf{x}_{it})$. This formulation shows that the Markov chain can 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). The third stochastic element of the HMM are the observation (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.$b_{it}^k(w) = P(O_{it} = w | Z_{it} = k, \mathbf{x}_{it})$. 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).

Data on illegal cartels

In modern data sets on discovered cartels (see, e.g., Miller 2009, Brenner 2009, Levenstein Suslow 2006), the observed data vary but are becoming increasingly detailed. To illustrate such data, consider how 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. 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 (non-) existence in a given period. This observation process may produce 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. 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.

Table B1 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" $= c$, "Detected and shut down by the CA" $= d$, "Leniency" $= l$ and "Unknown / unobserved" $= u$.

Table B1: 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, $\mathbf{o}_1 = (u, u, u, d, u)'$. 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. The cartel investigations 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. 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 can be similarly interpreted; it enters the data in a cartel.

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 might correspond to an industry that was suspected and investigated for having a cartel over a two-year period. The records (e.g., the court decision) show 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.

A great advantage of a HMM is that it can easily be tailored to the specifics of the institutional environment. To show how, we outline next a HMM for illegal cartels that allows for a probability of cartel detection, and for a probability of applying for leniency, as in Harrington and Chang (2009) (see also Harrington and Chang 2015; henceforth jointly HC). These two probabilities are empirically important because they are key (structural) parameters describing the efficacy of modern competition policy. With further structure from HC one could also estimate the harm caused by the cartels.

Hidden process with illegal cartels

It is convenient to use slightly different parametrization of the hidden process for our purposes here. Therefore, and in spirit of HC, we let κ_{it} denote the probability that there is an opportunity in market i at time $t > 1$ to form a cartel (if there is none at $t-1$) and H_{it} denote the probability that a cartel is stable (i.e., that the incentive compatibility constraint, ICC, holds). We also assume that 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 potential 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. Following HC we postulate that firms resort to the leniency program only if the cartel is breaking up.[1] Conditional on it happening, the probability that the cartel 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 the hidden process for Z_{it}, its state space is $S_Z = (n, c, d, l)$. The associated transition matrix $\mathbf{A}_{it}$ is

[1] In practice, the probability of a leniency application may be a function of cartel detection/leniency in related markets. Variables capturing such events could be introduced into the empirical model as state variables.

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

The elements of $\mathbf{A}_{it}$ are the transition probabilities of a first-order 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, it may turn out that the cartel is not stable (i.e., the ICC does not hold), 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.

We have specified $\mathbf{A}_{it}$ with particular assumptions in mind. First, the detection probability σ_{it} shows up only in columns two and three because we assume that 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. Second, 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. Both of these assumptions can be relaxed if the institutional environment so requires and/or if the available cartel data are rich enough to permit a more flexible model (e.g. a larger state space). For example, one could allow for the possibility that detection affects subsequent re-formation of a cartel.

Observation process

In modern era data sets, the state space of the observation process is determined by the institutional environment and the available data. We augment the state space here to $S_O = (n, c, d, l, u)$, where "*d*" refers to a cartel that has been detected and shut down by the CA and "*l*" to a leniency application. This kind of observed data can be linked to the hidden process in many ways.

For example, assume that (i) if an industry is (is not) in a cartel, the observed data never wrongly suggest that it is not (is), that (ii) the exposure of a cartel to the CA is observed (by the researcher) with probability one, and that (iii) 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. The observation probability matrix would then be

$$\mathbf{B}_{it} = \left[\; b_{it}^{k}(w) \;\right] = \begin{bmatrix} \beta_{it}^{n} & 0 & 0 & 0 & 1-\beta_{it}^{n} \\ 0 & \beta_{it}^{c} & 0 & 0 & 1-\beta_{it}^{c} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

where $b_{it}^{k}(w)$ again denotes the probability of observing $w \in S_O = (n, c, d, l, u)$ when the unobserved state of industry i at time t is $k \in S_Z = (n, c, d, l)$, and $b_{it}^{n}(n) = \beta_{it}^{n}$ and $b_{it}^{c}(c) = \beta_{it}^{c}$. Parameters β_{it}^{c} and β_{it}^{n} reflect the ability of the CA (and courts) to determine whether a detected cartel did or did not exist in the periods prior to the detection. They are therefore potentially policy relevant.

Other assumptions about the observation process would lead to a different $\mathbf{B}_{it}$. For example, the assumption of no labeling mistakes could be relaxed. Provided that the data are rich enough, one could allow $b_{it}^{c}(n)$ and $b_{it}^{n}(c)$ to be nonzero to be in line with Lee and Porter (1984) and Ellison (1994).

Estimation and identification

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

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

where $\mathbf{o}$ denotes the data (the realization of $\mathbf{O}$). This shows that the elements of $\mathbf{D}_{it}$ needed for the likelihood can be derived from $\mathbf{A}_{it}$ and $\mathbf{B}_{it}$. There are 4+2 probabilities that call for identification in this HMM of illegal cartels. Identification is very similar to the model with legal cartels. The new parameters σ_{it} and ν_{it} are identified from transitions to and from d and to and from l.

So far, we have been agnostic about the precise form of H_{it}. If one wants to impose structure to it, the models of HC would give a good starting point. One can, for example, modify the ICC condition so that the model explicitly allows for CA detection and leniency. With data on illegal cartels, the returns to structural estimation of H_{it} are likely to be high, as it would allow a number of interesting counterfactual experiments on competition policy.

Appendix C: Estimation figures and robustness tests

Here we first present the figures for $H1$ and $H2$ from our baseline model (see Table 3 and Figure 2) with confidence intervals. We then present the $H1$, $H2$ -figure and the estimated degree of cartelization for the specification using a polynomial of time and for our robustness tests. In the figures, "Estimated" refers to the estimated degree of cartelization and, if shown, "Macro shocks smoothed" refers to the counterfactual calculation of cartelization in which the largest positive GDP shocks have been reset to the mean value of the shocks in the sample. Finally, we report the AIC-weights used to produce the model-averaged version of the degree of cartelization displayed in Figure 3.

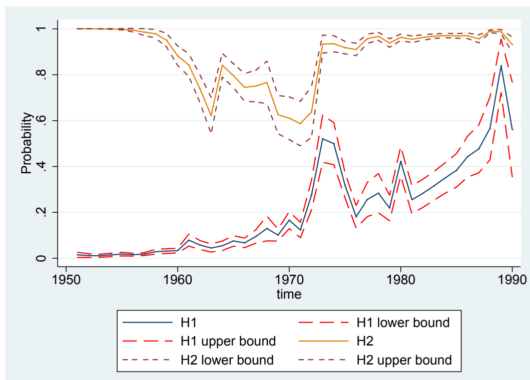

Figure C1: $H1$ and $H2$ from the baseline model with 95 percent confidence intervals

Figure C2: $H1$ and $H2$ and cartelization using polynomial of time (no macroshock in the model)

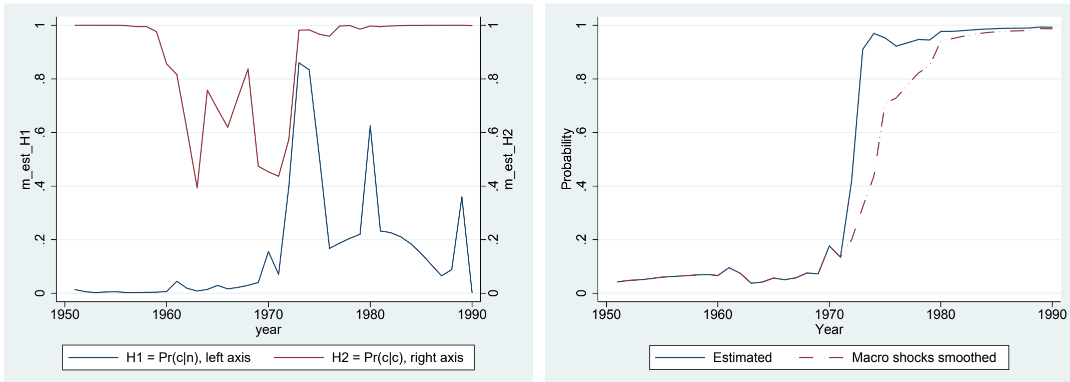

Figure C3: $H1$ and $H2$ and cartelization using one cartel markets

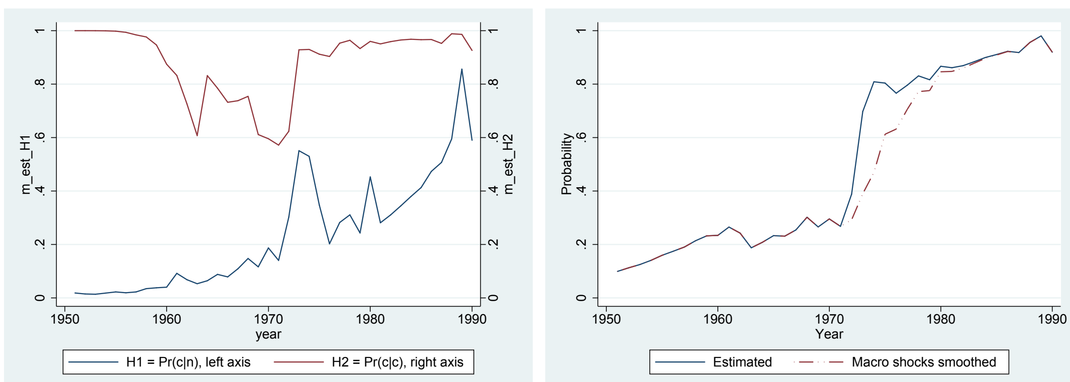

Figure C4: $H1$ and $H2$ and cartelization using data from 1959 onwards only

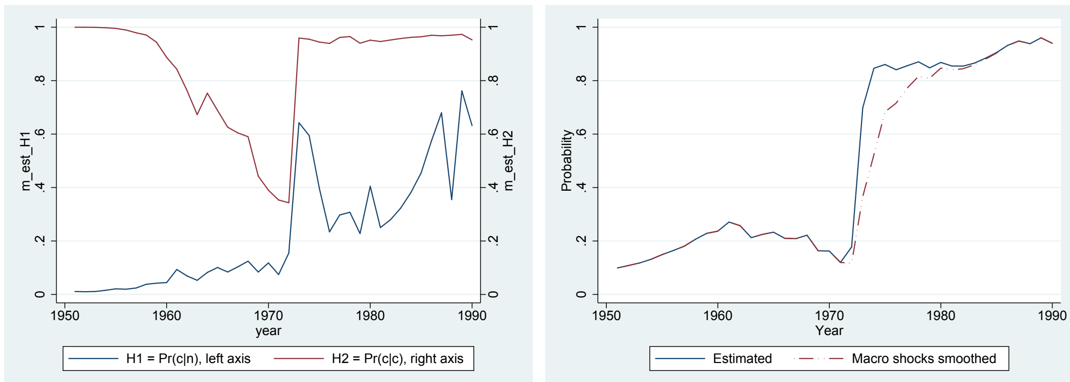

Figure C5: $H1$ and $H2$ and cartelization using law spline

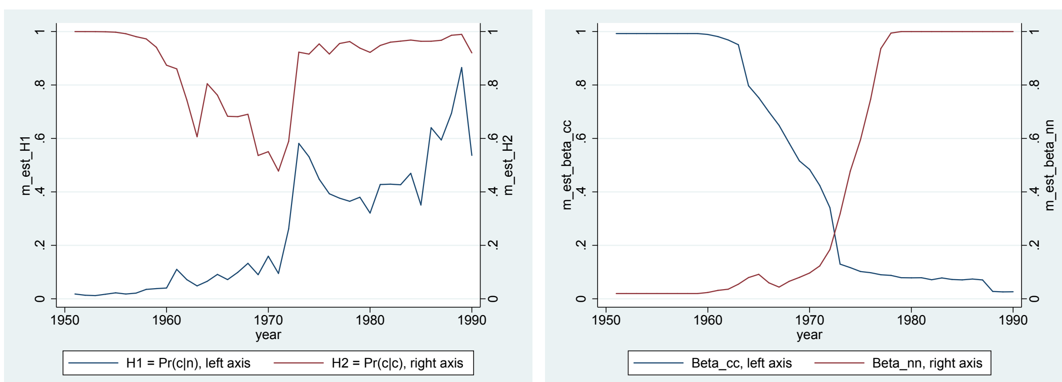

Figure C6: $H1$ and $H2$ and cartelization with observed heterogeneity

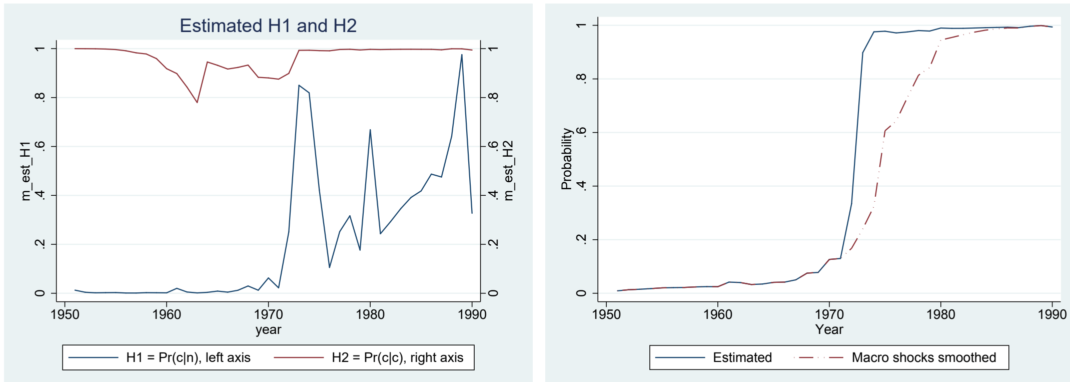

Figure C7: $H1$ and $H2$ and cartelization using 11 markets / industry

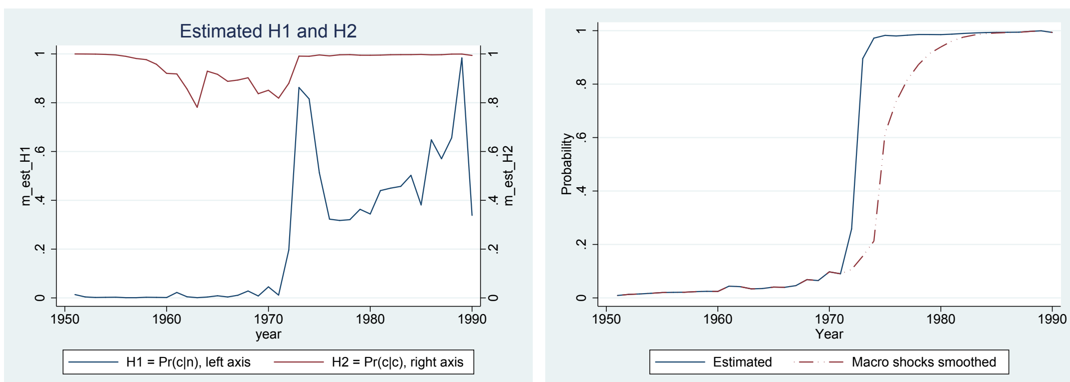

Figure C8: $H1$ and $H2$ and cartelization using 11 markets / industry & observed heterogeneity

Table C1: Weights in AIC model averaging

Model	Baseline	Time polyn.	Law spline	Law-index 85	Observed heterog.	Export
Robustness test #			3	3	4	5
Weight	2.47e-09	0	0.979	2.47e-09	0.005	0.016

Notes: AIC weights calculated according to Burnham and Anderson (2002).

Appendix D: Industry characteristics data 1974 - 1988

In this appendix, we describe the industry level variables that we utilize in two robustness tests (Section 5.5., Robustness test #4 and #7). The same variables are also used when we estimate the impact of cartelization on industry profitability (Section 5.6); this appendix also provides further details on these estimations.

Data source and variables

Our industry characteristics data come from Statistics Finland's Longitudinal Database on Plants in Finnish Manufacturing which is available to us for the time period 1974 - 1988. We use the available plant level data to generate industry level variables that are frequently mentioned as factors facilitating collusion. We use the following annual variables to model the hidden process:

Number of firms and concentration: The textbook supergame theoretic model of collusion suggests that collusion is harder to achieve, the larger the number of firms in the industry (e.g. Belleflamme and Peitz 2010 ch. 14.2). Similarly, it is commonly asserted (e.g. Carlton and Perloff 1990, pp. 221) that high concentration facilitates collusion. We measure concentration using the Herfindahl-index (HHI).

Asymmetry of firm size: Most of the theoretical literature suggests that asymmetry between firms makes collusion more difficult (e.g., Lambson 1994, Davidson and Deneckere 1984, 1990). Compte, Jenny and Rey (2002) find that this result depends on how large aggregate capacity is relative to demand. In most models of collusion, size and cost asymmetries make collusion harder (see e.g. the survey of Jacquemin and Slade 1989). We include the ratio of the sales of the second largest firm to the sales of the largest firm to capture the effects of asymmetry between the leading firms ($Ms - second - first$). One can argue that this variable also captures cost asymmetries.

Cost structure: The responsiveness of cartel prices to costs may vary, affecting incentives to collude (Harrington and Chen 2006). We include the ratio of material expenses to sales to measure the importance of variable costs ($Material - share$).

Product differentiation: The empirical literature suggests that collusion mostly occurs in homogenous goods industries (see e.g. Levenstein and Suslow 2006), but the theoretical literature addressing the same question portrays a more mixed picture. Chang (1991) and Ross (1992) find that differentiation makes collusion easier, while Raith (1996) and Häckner (1994) find the opposite. Thomadsen and Rhee (2007) show that costs of maintaining collusion increase the difficulty of sustaining collusion more for firms in industries with product differentiation. We allow for the potentially mixed effects of different types of product differentiation by including a dummy for the product of an industry being homogenous ($Homog - d$). This was constructed following the existing literature (Rauch 1999, Foster, Haltiwanger, Syverson 2008) by utilizing the characterization of each industry, and the Registry's description of the goods produced by the cartel (see also Hyytinen, Steen and Toivanen 2017).

Multimarket contact: Bernheim and Whinston's (1990) theoretical analysis shows that under certain conditions, such as cost asymmetries and scale economies, multimarket contact may facilitate collusion. The existing empirical research (e.g. Evans and Kessides 1994, Ciliberto and Williams 2013 and Molnar, Violi and Zhou 2013) provide evidence supporting this. We measure multimarket contact as the share of sales of the two largest firms in a given industry, calculated for industries where they are both present, excluding the industry for which we measure the variable ($Mm - share$).[2]

Industry growth: There is a large cartel literature focusing on the importance of demand fluctuations for cartels (see Levenstein and Suslow 2006 for a review). 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). 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 (Jacquemin, Nambu and Dewez 1981 and Suslow 2005). In addition to variables capturing the overall macroeconomic conditions, we also include industry growth to control for these effects ($Growth$), calculated as the annual growth rate of the gross value of deliveries.

Entry and exit: The lower the entry barriers, the more likely it is that a cartel that manages to raise prices invites more entry. We measure the ease of entry and exit by using the entry and exit rates of a given industry ($Entry - share$, $Exit - share$), calculated as the number of entering (exiting) firms divided by the existing stock of the firms in each industry.[3]

[2] The formula is the following: $Mm - share_{it} = \sum_{j \neq i} 1(sales_{ktjt} > 0) 1(sales_{mjt} > 0)(sales_{kjt} + sales_{mjt}) / \sum_{j \neq i} [1(sales_{ktjt} > 0) sales_{kjt} + 1(sales_{mjt} > 0) sales_{mjt}]$ where i, j index markets, t time, and k and m the largest and second largest firm in market i in year t. For two firms (1,2) present in two markets (A,B) and firm 2 also in market C, the equation reduces to e.g. $Mm - share_{At} = (sales_{1Bt} + sales_{2Bt}) / (sales_{1Bt} + sales_{2Bt} + sales_{2Ct})$. This formulation captures the importance of market B where both firms are present, relative to how important markets B and C in total are to the two firms.

[3] An alternative would be to calculate a churn-rate for each industry. In our view, the

Exports: While export cartels were not registered, they were both legal and in frequent use. The higher is the share of exports, the likelier it is that there is an export cartel in the industry, potentially facilitating cartelization also in the domestic market (Schultz 2002). We capture this by including the ratio of exports to the gross value of deliveries ($Export - share$).

Turnover: Finally, we include the industry level turnover, measured by the gross value of deliveries, to capture the effects of market size on cartelization ($Turnover$).

In the robustness tests reported in the main text, we allow all these variables to affect both the formation ($H1_{it}$) and the continuation ($H2_{it}$) probability. Despite these data being available only for 1974-1988, we use the full sample from 1951 to 1990 when estimating the HMM that includes the industry characteristics. In these estimations, the industry characteritics take value of zero for the years over which they are not available.

Estimating the effect of cartelization on profitability

To estimate the impact of cartelization on industry profitability, we proceed as follows: Our measure of cartelization is the industry-level price cost margin calculated as value added divided by turnover. Our measure of cartelization is the predicted probability of an industry having a cartel in a given year (using the recursive algorithm based on our Markov model), based on the estimation of our HMM where we use industry characteristics (see Section 5.5, Robustness test #4).

When estimating the impact of cartelization on industry profitability, our estimation sample, determined by the available data, is 193 industries for 14 years (13 for 3 industries), 1974 - 1987, yielding 2 699 observations. The annual time trend variable takes value one in 1974. We display the results in Table D1 (with robust standard errors in parentheses). The estimated model includes the annual time trend and industry fixed effects. The table shows that the likelihood of cartelization is positively associated with the price cost margin.

Table D1: Price cost margin and cartelization

$P(cartel)$	0.333***
	(0.054)
$time$	0.014***
	(0.001)
R^2	0.705
Industry FE	YES
N	2699
#industries	193
Notes: * p<0.10, ** p<0.05, *** p<0.01	

effects of entry and exit on collusion are potentially different, warranting the inclusion of both in the specification.

Appendix E: Assigning cartels to markets in case of multiple registered cartels

In the main body of the paper we treated each industry as an individual market. We report here an alternative way of constructing the data for our dependent varibale. This alternative procedure deals with the issue of observing more than one (simultaneous) cartel in a given industry in a more elaborate way by creating markets within each industry. Our main and alternative processes consist of three steps each, with both processes sharing the first two (described in the main text). In this appendix, we explain how the alternative procedure deals with those cases with multiple cartels in the same industry (i.e., step 3 of the alternative procedure). We also report the associated transition matrix and briefly mention what we find when we replicate our baseline HMM using the alternative data.

Assigning cartels to markets within an industry

There are two reasons for us observing more than one registered cartel in a given industry. The first reason is that an individual entry into the Registry (a "registered cartel") does not necessarily correspond to the economic definition of a cartel ("actual cartel"). In some cases, two registered cartels were clearly part of the same actual cartel. As an example, we compared the members of the registered cartels if they were assigned to the same industry by the Registry. If the members were the same and the purpose of the registered cartels interlined, we concluded them to be part of the same actual cartel.

The second reason, which we faced in the remaining industries, is that some registered cartels that operated in the same industry were clearly different entities. This became clear when comparing the verbal descriptions of some of the cartels assigned to the same industry. This is easy to deal with when the cartels are not overlapping timewise; we then simply assign them to the same industry (and market(s)), and view them as two or more consecutive cartels.

To deal with those cases where the cartels existed simultaneously in a given industry, we assume, consistent with most theoretical models, that there is at most one actual cartel in a given market at any point in time. We treat each industry as consisting of an exogenously determined number of markets.[4] We then assign each overlapping cartel to a separate market within the same industry. An outcome of this process is that we assign the value u for all years for those markets in a given industry for which there is no cartel.

An issue here is that some registered cartels are assigned to the same industry and were not part of the same actual cartel. To deal with this, we determined whether registered cartels assigned to the same industry are in the same market

[4] We need to assign at least as many markets to an industry as there are cartels. The maximum number of actual cartels per industry is 7. We chose the number of markets per industry to be 11, yielding us 2123 markets (as we have 193 6-digit industries). This assumption has no effect on our estimates.

and whether they are part of the same actual cartel, using qualitative information obtained from the Registry. The evidence consisted of the assignment of the registered cartels to SIC industries by the FCA, the qualitative description of the competition restriction by the FCA, and lists of members of the registered cartels. We then applied the following rules to industries with multiple registered cartels that were not part of the same actual cartel:

1. Those multiple registered cartels that were judged to be in different markets while in the same industry were each assigned to a separate market within the industry.
2. If the multiple registered cartels could not be assigned to different markets based on qualitative information but were sequential,[5] they were assigned to the same market.
3. If the multiple registered cartels could not be assigned to different markets based on qualitative information and were simultaneous, we assigned them to different markets, as suggested by most theoretical models.

As to Rule 2, we proceed by first coding the observed states for all cartels in the same market separately. We then merge these as follows: If we observe "c" for one cartel but "n" or "u" for the others in a given year, we assign "c" to that year on the basis that we know that at least one of the cartels was active in that year. If we observe "u" for one and "n" for some of the others in a given year, we assign "u" to that year on the basis that while we know that one of the cartels did not exist in that year, we don't know the status of the others. Rule 3 stems from the identification of our model which requires us to have at most one cartel at a given point in time in a given market.

Descriptive statistics with multiple markets per industry

The alternative procedure results in a dataset where we have 40 annual observations (1951-1990) for 2123 markets in 193 industries, yielding 84920 market-year observations. Table E1 shows the transition matrix of our dependent variable, using the alternative data. We have 360 observations for which we know for consecutive years that a cartel did not exist in a given market in either year. Similarly, we observe 641 cases where a cartel existed in two consecutive years. As can be seen, the vast majority of transitions are between two consecutive market-year observations where we do not know whether a cartel existed or not. All in all, the u- observations account for 98percent of the data; this high number corresponds with the observation that most of the time, we don't know whether or not a given market has a cartel. In our data, this is partly due to the fact that if no cartel in the Registry is assigned to a given industry, all market-year observations in the industry are assigned u.

[5]We use information on the real and registry formation and continuation of cartels to determine whether they are simultaneous or sequential.

Table E1: Transition matrix

$t-1$ / t	n	c	u	total
n	360	113	142	615
	58.54%	18.37%	23.09%	100.00%
c	85	641	319	1 045
	8.13%	61.34%	30.53%	100.00%
u	180	272	80 685	81 137
	0.22%	0.34%	99.44%	100.00%
total | count	625	1 026	81 146	82 797
total | %	0.75%	1.24%	98.01%	100.00%

Notes: The number of observations in the table is 2 123 less than the number of observations in the data, as the transition cannot be calculated for the first year of the data.

Estimation results using the alternative data

When we replicate our baseline HMM using the alternative data, $H1$ increases trend-like and $H2$ remains high (see Section 5.5, robustness test #7, and Appendix C). The degree of cartelization implied by these estimates is very similar to what we report in the main text (see Appendix C). In an earlier version of our paper, we implemented a number of robustness tests using these alternative data. These robustness tests showed that changing the number of exogenously determined markets within an industry, using only industries with one observed cartel, and using data only on the first registered cartel in a given industry had no effect on our results. These results were as expected, as increasing the number of markets only leads to a higher fraction of observations in the (u, u) - cell of the transition matrix of the observation process. Such observations do not contribute to identification.

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] Belleflamme, Paul, and Peitz, Martin, 2010, *Industrial Organization: Markets and Strategies*, Cambridge University Press.

[3] Bernheim, Douglas, and Whinston, Michael, 1990, Multimarket Contact and Collusive Behavior, *RAND Journal of Economics*, Vol. 16, pp. 269-281.

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

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

[6] Carlton, Dennis, and Perloff, Jeffrey, 2004, *Modern Industrial Organization*, 4^{th}edition, Prentice Hall.

[7] Chang, Myong-Hun, 1991, The Effects of Product Differentiation on Collusive Pricing, *International Journal of Industrial Organization*, Vol. 9, pp. 453-469.

[8] Ciliberto , Federico and Williams, Jonathan, 2013, Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conduct Paratemers in the Airline Industry, CEPR DP9015.

[9] Compte, Olivier, Jenny, Frederic, and Rey Patrick, 2002, Capacity Constraints, Mergers, and Collusion, *European Economic Review*, Vol. 36, No. 1, pp. 1 - 29.

[10] Davidson, Carl, Deneckere, Raymond, 1984, Horizontal Mergers and Collusive Behavior, *International Journal of Industrial Organization*, Vol. 2, pp. 117-132.

[11] Davidson, Carl, Deneckere, Raymond, 1990, Excess Capacity and Collusion, *International Economic Review*, Vol. 31, pp. 521-541.

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

[13] Evans, Willian and Kessides, Ioannis, 1994, "Living by the Golden Rule": Multimarket Contact in the U.S. Airline Industry, *Quarterly Journal of Economics*, Vol109,

[14] Foster, Lucia, Haltiwanger, John and Syverson, Chad, 2008, Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, Vol. 98. No. 1, pp. 394-425.

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

[16] 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.

[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] Harrington, Joseph E., Jr. and Chang, Myong-Hun, 2015, When Should We Expect a Corporate Leniency Program to Results in Fewer Cartels?, *Journal of Law and Economics*, Vol. 58, Vol. 2, pp. 417-449.

[19] Harrington, Joseph E., Jr. and Chen, Joe, 2006, Cartel Pricing Dynamics with Cost Variability and Endogenous Buyer Detection, *International Journal of Industrial Organization*, Vol. 24, No. 6, pp. 1185-1212.

[20] Häckner, Jonas, 1994, Collusive Pricing in Markets for Vertically Differentiated Products, *International Journal of Industrial Organization*, Vol. 12, pp. 155-177.

[21] Hyytinen, Ari, Steen, Frode, and Toivanen, Otto, 2017, Anatomy of Cartel Contracts, manuscript, Aalto University, available from: http://aalto-econ.fi/toivanen/.

[22] Jacquemin, Alexis, and Slade, Margaret, 1989, Cartels, Collusion and Horizontal Mergers, in *Handbook of Industrial Organization* Vol. 1, Richard Schmalensee (ed.), North-Holland.

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

[24] Lambson, Val E., 1994, Some Results on Optimal Penal Codes in Asymmetric Bertrand Supergames, *Journal of Economic Theory*, Vol. 62, pp. 444-468.

[25] 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.

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

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

[28] Molnar, Jozsef, Violi, Roberto and Zhou, Xiaolan, 2013, Multimarket Contact in Italian Retail Banking: Competition and Welfare, *International Journal of Industrial Organization*, vol. 31, No. 5, pp. 368 - 381.

[29] Raith, Michael, 1996, Product Differentiation, Uncertainty and the Stability of Collusion, Discussion paper EI/16, The Suntory and Toyota International Centres for Economics and Related Disci-plines, LSE

[30] Rauch, James E., 1999, Networks versus Markets in International Trade, *Journal of International Economics*, Vol. 48, pp. 7–35.

[31] Ross, Thomas W., 1992, Cartel Stability and Product Differentiation, *International Journal of Industrial Organization*, Vol 10, pp.1-13.

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

[33] Schultz, Christian, 2002, Export Cartels and Domestic Markets, *Journal of Industry, Competition and Trade,* Vol. 2, No. 3, pp. 223-246.

[34] 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.

[35] Thomadsen, Raphael and Rhee, Ki-Eun, 2007, Costly Collusion in Differentiated Industries, *Marketing Science*, Vol. 26 No. 5, pp. 660-665.

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