The effect of the Sarbanes-Oxley Act on industrial innovation – evidence from patenting of stem cell technologies

Waters, James

University of Nottingham

9 June 2013

The effect of the Sarbanes-Oxley Act on industrial innovation - evidence from patenting of stem cell technologies

James Waters [*][†]
Nottingham University Business School

Working paper - comments welcome

June 9, 2013

Abstract

This paper assesses how financial regulation, in the form of the Sarbanes-Oxley Act (SOX), affects industrial innovation. We describe the decision faced by an investor considering an innovative project before, at the time of, and subsequent to SOX. SOX is shown to create a real option value for investment delay through its short term hike and medium term diminution in uncertainty, and we calculate how implementation parameters impact on delay and subsequent innovation. To test our model's implications, we examine patenting in stem cell technologies from 2001 to 2009, introducing two methods of endogenously detecting changes in patenting rates. We find a post-SOX dip in patenting consistent with our model, and no long term effect on innovation. We reject US government funding cuts to human embryonic stem cell research as a cause of the observed behaviour, as well as five other sets of explanations.

1 Introduction

The US Sarbanes-Oxley Act (Sarbanes-Oxley, 2002, hereafter SOX) was a regulatory response to financial scandals in the early 2000s. It imposed new disclosure requirements on companies, placed increased restrictions on the behaviour and conflicts of interests for company insiders, and specified new penalties for related malfeasance. Auditors and securities analysts were also subject to more stringent regulations on their behaviour, and a

---

[*] University of Nottingham Business School. Correspondence: Nottingham University Business School, Jubilee Campus, Nottingham, NG8 1BB, United Kingdom. E-mail: james.waters@nottingham.ac.uk. Part of the work on this paper was done while the author was a visiting lecturer at the University of Westminster Law School.

[†] The author thanks Pelin Demirel and Joe Tanega for helpful discussions, and seminar participants at Nottingham University Business School. All mistakes are my own.

new not-for-profit company, the Public Company Accounting Oversight Board (PCAOB), was established to oversee auditors and audit quality.

Many regulatory demands of SOX could incur costs for companies. The PCAOB, for instance, is funded by company levies, and the increased demands on auditor quality could be passed on as greater charges for audit. The net effect of SOX on available company funds is not apparent as the Act may increase investor confidence in investment at the same time as increasing the effective charges levied against that investment. Neither is it clear what the change will be in incentives for particular managerial uses of the available funds. For example, activities that are more heavily regulated in the new rules may be relatively less profitable relative to other activities after SOX.

Given these uncertain effects on company funding and incentives, the literature has examined the association between the passage of SOX and contemporaneous changes in a variety of financial and non-financial corporate variables. They include market return (Li ***et al.***, 2008; Zhang, 2007), equity cost (Ashbaugh-Skaife ***et al.***, 2009), investment discount factors (Kang ***et al.***, 2010), cross-listing premium (Litvak, 2007), company cost and structure (Linck ***et al.***, 2009), market entry choices (Engel ***et al.***, 2007; Leuz ***et al.***, 2008), profitability (Ahmed ***et al.***, 2010), investment efficiency (Cheng ***et al.***, 2013), and research and development spending (Bargeron ***et al.***, 2010). Some research has examined changes in variables indicating mechanisms by which SOX may affect the market, including executive incentives (Cohen ***et al.***, 2013), costs (Millar and Bowen, 2011), informational asymmetry (Akhigbe ***et al.***, 2010), informational content of earnings (Singer and You, 2011; Kalelkar and Nwaeze, 2011), and investment risk (Akhigbe ***et al.***, 2008; Akhigbe and Martin, 2008).

This paper examines whether SOX was associated with a change in industrial innovation by affected companies. SOX may have altered their available funding for research and development or managerial preferences for such risky activities as found by Bargeron ***et al.*** (2010) and Cohen ***et al.*** (2013). Innovation represents an investment in intellectual property made in anticipation of a stream of future income from its commercialisation, and the decision to invest depends on revenues taking into account SOX's effects on costs and risk.

We formulate the dynamic problem faced by an investor considering investment in an innovative company. Managerial quality is unknown, which introduces uncertainty in returns. Prior to SOX, the investor's problem is time symmetric and there is no benefit from waiting. When SOX is initially introduced, additional uncertainty is introduced linked to SOX's costs. The uncertainty about SOX's costs and managerial quality declines over time, and the investor is incentivised to delay investment both by the real option value of waiting and the decline in the managerial quality risk premium. In the long run, SOX's effect on investment depends on the relative sizes of SOX costs and the risk premium. We test for innovation's temporary fall after SOX's introduction, and infer SOX's net effects independent of temporary uncertainty effects by comparison of long run innovation with pre-SOX innovation.

Our testing and inference is on the time series of patent rates in stem cell technologies. Our reasons for selecting stem cell technologies is fourfold. The biotech industry has itself raised concerns about the effects of SOX (Biotechnology Industry Organization, 2006). The stem cell technology industry is new, small, and risky. These characteristics have often been found to be linked to magnification of SOX's effects on funding and incentives (Chhaochharia and Grinstein, 2007; Engel ***et al.***, 2007; Jain ***et al.***, 2008; Kamar ***et al.***, 2007; Kang ***et al.***, 2010; Li ***et al.***, 2008; Litvak, 2007; Wintoki, 2007). It is an industry that is of particular social prominence and potential. Finally, selection of a single industry for examination gives data amenable to our estimation techniques, which are univariate in order to facilitate flexible, endogenous break date estimation.

SOX's effect can be identified by comparing quantities of interest in base periods assumed to have been influenced by SOX with their value in control periods assumed to be free of influence. Existing literature differs in the length and timing of the base and control periods. Event studies such as Li ***et al.*** (2008) and Zhang (2007) take short base periods associated with plausibly significant SOX-related occurrences and announcements, and usually extended control periods around the base periods. Doidge ***et al.*** (2009) use a long control and base period with a single cut-off point at the end of 2001 (as well as a more flexible sequence of annual cross sectional regressions), Cohen ***et al.*** (2013) use the same cut-off point, while Ahmed ***et al.*** (2010) have a control period ending in 2002 and a base period starting in 2004. Engel ***et al.*** (2007) use both short and long base periods to classify times by SOX's impact. These periods are generally ultimately set according to the researchers' judgement.

It is not clear which dates to include in each period in view of SOX's decade long planning, passage, implementation, delay, and revision (Gao ***et al.***, 2009; Biotechnology Industry Organization, 2012). The theoretical case for including, say, the date on which a senator issued a statement of regulatory intent, may be uncertain. Further, the effect may spill beyond the period hypothesised, or it may be delayed, or it may be a temporary reaction that is subsequently corrected. Alternative date classifications can lead to different conclusions. These classification problems are particularly severe for short periods. Whilst longer periods mitigate the difficulty, they are more likely to have influential non-SOX events within them, leading to problems in ascribing any changes to SOX.

Our approach to the classification problem is to use the data to estimate change dates in the number of US residents' patent applications over the period 2001 to 2009. Our first procedure estimates an autoregressive negative binomial count process in the presence of a structural change in trend and level where the change date varies across the whole period, leading to a series of change p-values. We examine the p-values to identify the absolute and relative importance of any break that occurred around the time of SOX. We use direct comparison of the p-values in the eighteen month period from July 2002 to December 2003 to identify the date at which a break was most likely. The coefficient estimates at that date are reported, as are Monte Carlo simulated p-values for the break coefficients. These values correct for the conditionality introduced in the raw p-values by the search procedure, and for the bias induced by small sample maximum likelihood estimation.

Our second procedure allows for multiple breaks and multiple patenting states. To identify them, we assume a constant probability of transition between Poisson states, with possibility of re-entry to previously exited states. The resulting hidden Markov model is estimated to give the most likely series of states over the whole period, and the coefficient estimates within each state. The estimation is performed for two and three distinct states.

The p-value testing finds that structural change occurred with higher probability around the time of SOX than elsewhere in the period. A trend change in patenting was most likely around July 2002, while a level shift was most likely around July 2003. In the two state hidden Markov model, a pre-SOX change is identified where patenting starts to increase from a very low level. The three state hidden Markov model identifies another change in July 2003, when the rate of patenting declines. Consistent with the earlier theoretical predictions following from the modelling assumptions, the rate corrects after two years to its previous process.

The timing of the patenting process break is consistent with other feasible non-SOX explanations. They include shocks to the patenting process that affect US and non-US patenters equally (unlike SOX), shocks that affect institutional and non-institutional patenters equally (unlike SOX), an exhaustion of the technological prospects of US patenters, a recent industry downturn that had driven many US innovators out of the market, reduction in US government funding for stem cell research, and shocks that affect only stem cell innovation (unlike SOX). We consider their plausibility and some empirical hypotheses that are closely related to them. Evidence is presented against them as exclusive explanations for the observed break.

Our work is related to a number of research strands in the literature. It contributes to the analysis of how financial changes affect industrial outcomes. The impact of finance on innovation is well recognised Hall (2002), and there have been some studies of how changes brought by financial crises affect innovation (Paunov, 2012; Archibugi ***et al.***, 2013). Bargeron ***et al.*** (2010) look at how SOX affected R&D. However, this study is the first we are aware of that looks at how innovative outputs are affected by financial regulation and specifically SOX.

For the general SOX literature, we believe that this paper represents the first real options representation of investment decisions after the Act. It indicates that the common examination in the financial literature of SOX effects over a period of several years is likely to be misleading because of uncertainty and risk option effects. We believe we are the first to employ hidden Markov models as an econometric method able to disclose corrections in post-SOX outcomes, and the first to endogenise break dates, through HMMs and the introduction of a robust time-varying break methodology.

Section 2 looks at our theoretical framework, section 3 describes our data, the empirical method is described in 4, results are presented in section 5, and section 6 concludes.

2 Theoretical framework

In this section we describe the options faced by an investor who can invest in an project leading to innovation. Their investment criteria are given prior to SOX, immediately following its introduction, and in the long run. After solving the criteria, we use them to formulate a testable hypothesis and a means of inferring SOX's persistent effect.

An investor is profit maximising and earns income from investing in a company and receiving profits from their investment. They discount future income at a rate d. The investor is risk averse and incurs costs of K_R when they receive a random income drawn from distribution function R; this cost may arise from extra capital being advanced to protect against loss exposure and so not being available for higher yielding projects elsewhere. When R is concentrated at a single point, $K_R = 0$. Investors have funds available for immediate investment in the company, or can choose to delay for m time periods and then reassess their options.

The company requires investment of C. If the company receives the investment, it is spent on an innovative project that is non-reversible and there are no resale opportunities at a value above the investment return. Innovation results in intellectual property and a one-off income from it of r_t. Because of difficulties in observing the company's managerial quality, the investor perceives r_t as unknown prior to investment and treats it as a random variable drawn from a distribution R_t. The company protects its intellectual property by patenting, with the number of patents increasing in r_t.

Prior to SOX, the distribution R_t is unchanging over time, so R_0= R_m. When SOX is introduced, the income generated after investment decreases by a one-off cost of s_t. We call this extra term "cost" by reference to the additional auditing and other expenses incurred through SOX, but it may be negative and so increase the investor's income if the changes improve the ongoing performance of the company. Investors are initially unsure about the size of s_t, and treat it as drawn from a distribution S_t. They have assumptions about how they will update their prior distribution of the effect as information emerges over time. Companies anticipate that they will have sufficient information to form stable judgements of SOX's effect by time m, so that the support for S_m will be a single point within the support for S_0. We denote by Q_t the distribution of the sum $q_t = r_t - s_t$.

Before SOX's introduction, the investor loses from delay in making their investment decision or not. The distribution function of returns is stable over time, so the same undiscounted expected return will be earned from immediate or delayed investment, and the latter return is subject to discounting. So investment will be made if and only if the expected returns from immediate investment net of risk costs are positive, or

$$\int_{I(R_0)} r \; dR_0 - K(R_0) > 0 \tag{1}$$

where $I(R_0)$ is the support of R_0.

Immediately after SOX's introduction, the returns from investment change over time so a choice not to invest immediately does not imply a permanent commitment. Thus, investors choose whether to invest straightaway or delay for m periods until they are fully informed about SOX costs and managerial quality. Delay occurs if and only if

$$\int_{I(Q_0)} q \; dQ_0 - K(Q_0) < \frac{1}{(1+d)^m} \int_{I(Q+)} q \; dQ_0 \tag{2}$$

where $I(Q_0)$ is the support of Q_0 and $I(Q+) = \{q \in I(Q_0) : q > 0\}$

The left hand side is the expected value from immediate investment, and the right hand side is the expected value from waiting and only investing if the revealed managerial quality net of costs gives a positive profit. We can decompose the integral on the left hand side to give

$$\int_{I(Q+)} q \; dQ_0 + \int_{I(Q-)} q \; dQ_0 - K(Q_0) < \frac{1}{(1+d)^m} \int_{I(Q+)} q \; dQ_0 \tag{3}$$

where $I(Q-) = I(Q_0) \setminus I(Q+)$, which is the set of SOX costs and managerial qualities that result in a loss if they occur after m periods.

Thus, the criterion for delay is

$$1 - \frac{1}{(1+d)^m} < \frac{- \int_{I(Q-)} q \; dQ_0 + K(Q_0)}{\int_{I(Q+)} q \; dQ_0} \tag{4}$$

We term the right hand side minus the left hand side the value of delay; we could alternatively introduce an error term into a stochastic investment decision criterion and frame our next results in terms of investment probabilities. Inequality 4 can be used to assess the impact on short run investment delay of features of SOX. For example, with parameter values of $m = 1$ and $d = 10\%$, the left hand side equals 0.09. The real option value of waiting is so substantial that a potential loss from ignorance or a risk aversion that are quite small relative to the potential profit from early investment will lead to delay.

If the expected costs of SOX increase, then the integral in the numerator rises and the integral in the denominator falls. The right hand side becomes larger and the value of delay increases. Intuitively, there is less to be gained from early commitment and more to be lost.

If the dispersion in the expected costs increases without altering the expected value, then any extra expected profits will be exactly offset by extra losses and they will grow monotonically with the dispersion (we recall that SOX costs may be positive or negative if there are managerial improvements that follow from it). Thus, the absolute values of the integrals in the numerator and denominator will increase and the value of the right hand side of the inequality will tend to unity monotonically. If the starting value of the numerator exceeds the starting value of the denominator then delay initially occurs. The convergence is from above and since the left hand side of the expression is strictly less than one, delay will continue to occur no matter how much dispersion there is. On the other hand, if the starting value of the denominator exceeds the numerator it is possible that the investor initially prefers to invest immediately, and the increased dispersion can lead to delay as the value of the right hand side increases. Thus, increased dispersion can lead to investment delays, which is reasonable as increased dispersion can reduce the relative importance of expected gains from early investment (that is, increase the right hand side ratio from below unity) and so make delay preferable.

As the speed of uncertainty reduction increases, then the parameter m reduces on the left hand side of the inequality 4. The value of the left hand side falls, and delays are more likely. The intuitive explanation is that if there are less discounting costs incurred by waiting to see what the costs and managerial quality are, then the value of delay increases.

From the considerations on SOX costs, delay, and uncertainty reduction, we can infer how SOX implementation might lead to delay. As SOX certainly introduced uncertainty about its costs, with the promise of reduced uncertainty over time about them and about the level of managerial quality, increased delays would be expected. The effect of net SOX costs is ambiguous, as it is unclear whether they are positive or negative, although the literature has reported net positive costs. On balance, we will test the following hypothesis.

Hypothesis 1. A temporary downward shift in innovative investment will occur after SOX.

We can strengthen the hypothesis slightly. Hartman (2007) presents data from 2001 to 2006 on SOX related costs and companies' self-reported ability to predict them. The survey results show (pages 16-7) stabilisation of overall costs of operating as a public company after 2004, for companies with annual revenue over and under $1 billion. The stabilisation is attributed to offsetting internal efficiencies, as actual external payments such as audit fees continue to increase (pages 2 and 19). They further show (page 14) a large increase between 2004 and 2006 in companies agreeing or strongly agreeing with the statement "...I am better able to predict costs associated with corporate governance reforms", from 38% to 65%. These results do not translate to an exact parameterisation for our model. However, they indicate narrowing distributions close to a point or slow moving deterministic process if the prediction process for future distributions has additional plausible assumptions, such as more weight being given to recent observations than older ones, importance being assigned to small changes in recent costs, and increased confidence in prediction resulting in diminishing importance of wide prior distributions. Thus, we may expect most learning about SOX to happen over a period of a few years, so that the restoration of patenting after SOX should happen on the same timescale.

After m periods, investors are aware of SOX's costs and managerial quality and face no further uncertainty. They therefore face a problem of investing if and only if

$$r_m - s_m > 0 \tag{5}$$

where r_m and s_m are realisations from the returns distribution R_0 and the SOX cost distribution S_0 respectively.

Comparing with the investment criterion before SOX given in inequality 1, we see that the long run value of investment will increase if and only if

$$r_m - s_m > \int_{I(R_0)} r \; dR_0 - K(R_0) \tag{6}$$

which says that the observed returns net of the observed SOX costs exceed the expected returns net of the risk premium. Generally the left hand side and right hand side will differ in value, and SOX will induce a long term change in investment and innovation. At the time of SOX's implementation the two left hand values will not be known, and their expected values can be taken to determine whether SOX is expected to increase investment. Cancelling out the terms for expected returns and rearranging, we have

$$K(R_0) > \int_{I(S_0)} s \; dS_0 \tag{7}$$

where $I(S_0)$ is the support of S_0.

So SOX is expected to increase investment if and only if the risk premium from ignorance about managerial quality (which vanishes after SOX) exceeds the expected SOX costs (which appear after SOX). If investors are not risk averse and SOX is expected to be costly, then the left hand side will be zero and the right hand side will be positive, so the value of investment will tend to fall after SOX.

With our available data, we will not be able to test these theoretical results in a formal hypothesis. However, we will interpret changes in our analysed quantities as being due to SOX, in line with the assumption in the literature and subject to exclusion of alternative causes. We differ from the literature in the timing of the comparison; rather than comparing immediately after SOX, we compare only after any dip in patenting has been corrected, as if our hypothesis is validated the initial period is characterised by adjustment to uncertainty rather than persistent cost and benefits.

3 Data

Our empirical analysis examines the patenting rates for innovation in the area of stem cell technology. The stem cell technology industry is recently established, with only 22 issued patents referring to stem cells in their title by the end of 1995 (USPTO, 2013), shortly following the establishment of current industry leaders such as Geron, Osiris, and Viacord. The industry remains relatively small, with Lysaght *et al.* (2008, Table 2) estimating 2007 commercial sales at $273 million, with average sales per operating company of $11 million. In the early 2000s, some of the industry's leading stocks exhibited considerable volatility (Salter, 2005). Investors at that time faced additional uncertainty relative to the situation today, in that the science was at a preliminary stage, product sales were negligible, and a number of the pioneering commercial enterprises in tissue engineering such as Organogenesis and Advanced Tissue Sciences had filed for Chapter 11 bankruptcy. Investment, company numbers, and sales were all rapidly increasing and volatile (Salter, 2005, Table 3; Salter, 2005, Table 1; Lysaght *et al.*, 2008, Figure 1). These characteristics indicate that companies innovating in stem cell technologies would be among those the literature finds to be most susceptible to SOX's impact. The impression is reinforced by reports from the US biotechnology industry body, the BIO, which has stated that "small biotechnology companies have been suffocated by the onus of complying with the Sarbanes-Oxley (SOX) Act" (Biotechnology Industry Organization, 2006), that "[t]he biotechnology industry was particularly hard-hit by the complex and burdensome regulations imposed by SOX's Section 404" (Biotechnology Industry Organization, 2007a), and that even after implementation reforms SOX has a "lack of clear direction to auditors" in respect of small company classification (Biotechnology Industry Organization, 2007b).

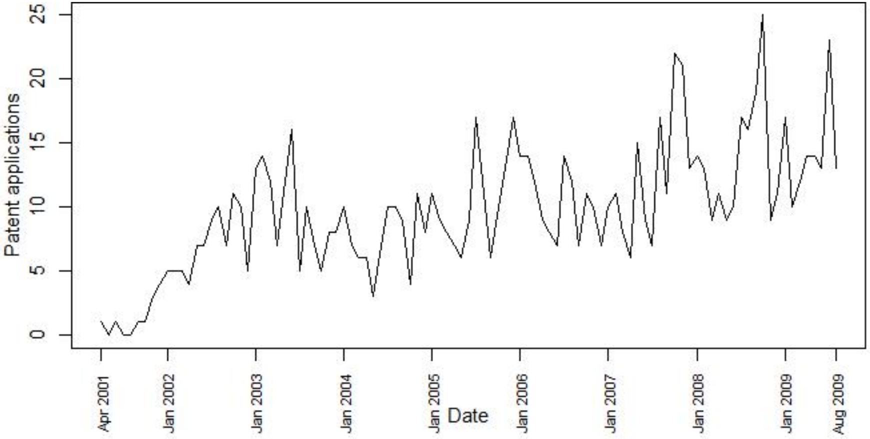
Our data on patents is from the US Patent and Trademark Office online database at USPTO (2013). We use patent applications as our measure of patents rather than granted patents because there is a delay between application and issuance while the patent examiner assesses whether the application meets the criteria for granting a patent. The date of application thus better measures the time at which the decision to commercialise was made. The USPTO records both the date of application and the date at which the application information was published on their database which is later than the date of application. Thus, some recent patent applications are omitted from the database, with more recent times having a smaller proportion of their applications recorded. To avoid the risk of a fall in applications due to delayed coding interfering with our attempts to identify changes in applications due to other factors, we identify patent applications with the date at which their information was recorded on the database. This date is a noisy and lagged measure of the application date. Table 1 shows the summary statistics for the distribution of times between filing and publication for stem cell technology applications published in 2003. Half of all patent application filings are published within 39 weeks, and three quarters are published in just over a year. Some publications take place shortly after application, so that an event affecting the decision to apply for patents should be reflected in changes in applications at contemporaneous publication dates with a near complete effect over the following two years.

Table 1: Summary statistics for the number of weeks between filing and publication of stem cell technology applications published in 2003

			Percentile			
Statistic	Mean	St.dev	50	75	90	95
Weeks	50	33	39	58	80	91

The USPTO database publishes applications only after March 1, 2001, with applicants after this date able to choose whether they wish to allow publication. Prior applications may not have been offered this choice, so would by default be deemed secret. Thus, most patent applications eligible for publication at much later dates would be made by applicants who could choose, whereas many patent applications eligible for publication at early dates may not have been published because the choice was not offered. This effect would be expected to lower the initial publication rate relative to later rates. Table 1 suggests that the effect will have substantially diminished by the middle of 2002 and all but vanished a year later. Its presence and timing complicates the interpretation of any change in the patent series as being due to SOX. We try to separate the two causes in section 5.2 by seeing whether the patenting of non-US applicants differed from that of US applicants, as both were subject to the USPTO regulations but generally only US applicants were subject to SOX.

Figure 1: Published patent applications by month for US residents



We used the USPTO online search facilities to identify patents with the exact phrase "stem cell" or "stem cells" in their title. The search was for published applications split by the month of publication for US resident applicants. Thus the code looked like

TTL/("stem cell" OR "stem cells") and PD/4/1/2001->4/31/2001 and ICN/("US").

The data covered the period from April 2001 to August 2009. 982 patent applications by US residents were published in the period. Figure 1 shows publications of patent applications by US residents. The rate of publications increases rapidly until around the middle of 2003, then steadies until around 2005 when it steps up to a new plateau. The plateau lasts until around 2007 when a new period of accelerating publication begins, lasting until the middle of 2009.

An alternative to searching by individual patents is to look for patent classes mentioning stem cell technology within their patent schedule or definition, and then count the number of patents within them. There are six such classes, listed in table 2. The classes do not exactly coincide with all stem cell patents, so patents not relating to stem cells are included and some stem cell patents may lie outside the classes. The classes can be narrowed down to subclasses to mitigate the latter problem, but may worsen the former.

Table 2: US Patent and Trademark Office patent classes with "stem cell" or "stem cells" in the description or schedule

Patent class	Description
424	Drug, bio-affecting and body treating compositions
435	Chemistry: molecular biology and microbiology
530	Chemistry: natural resins or derivatives; peptides or proteins; lignins or reaction products thereof
800	Multicellular living organisms and unmodified parts thereof and related processes
977	Nanotechnology

Generated by website searches on http://www.uspto.gov/web/patents/classification/ for "stem cell" and for "stem cells".

4 Empirical method

We present first the equations we will use in empirical estimation. They allow for the previous theoretical prediction of discontinuities in the patent process, through an autoregressive innovation metric combined with a current research input subject to breaks. We introduce two null hypotheses on the distribution of breaks in the patent process for testing an exceptional change around the time of SOX, and identify any changes with time-varying p-value and hidden Markov model estimations. Evidence for our earlier hypothesis can then be provided by finding downward shifts in patenting around the time of SOX's introduction, and subsequent correction towards the initial level within a few years. We can then infer SOX's net effects by comparing patenting after the correction and before SOX. Finally, we exclude a number of major potential alternative causes of the observed behaviour other than SOX.

Our base specification has published patent applications P_t at time t following a random process with conditional expected value at any time dependent on a numerical measure of investment conditions I_t at time t, exogenous research R_t until time t, and the history $H_t = \{P_{t-1}, P_{t-2}, P_{t-3}, \dots\}$ of past patent rates until time t. Expected patents are a free combination of a proportion of patents in each previous period and exogenous research. The specific form of conditional expected patent numbers is multiplicative and exponential in lagged patents and exogenous research. The multiplicative metric for innovation has been analysed in Blundell ***et al.*** (1995). The exponents depend on innovation activity, which itself depends on investment conditions. Given investment conditions, the exponents are constant and the consequent growth in patents captures both market entry by potential innovators and their innovation activity. Exogenous research grows at an exponential rate. Thus we have

$$E(P_t|I_t, R_t, H_t) = A(I_t)(e^{gt})^{B(I_t)} \prod_{i=1}^{\infty} P_{t-i}^{c_i(I_t)}$$

where g is a constant, and A, B, and c_i are functions.

Some of new potentially patentable innovations will have been anticipated by previous patents, and others will have been subject to accelerated approval by the patent office. These two effects are captured by dividing the expected patent numbers by fractional powers of lagged patents. The form of the equation is therefore unchanged. The exponents may be negative as well as positive.

Changes in investment conditions are represented by changes in the equation's coefficients. Given the analysis in section 2, conditions are variable at the time of the introduction of legislation and during the period of investors' learning about its effects.

Our first empirical specification has a single shock occurring over the period. We assume that the exponents on the lagged patents are constant before and after the shock, while allowing the drift and level to change. The testable equation is

$$E(P_t|I_t, R_t, H_t) = \exp(a + bt + d_1 DS_{t>u} + d_2 DT_{t>u}) \prod_{i=1}^{\infty} P_{t-i}^{c_i}$$

where $DS_{t>u}$ is a unit shift equal to one for times $t > u$ and zero otherwise, and $DT_{t>u}$

is a trend change defined by $DT_{t>u} = t - u$ for times $t > u$ and zero otherwise. The stochastic distribution is taken to be negative binomial, and the estimation is by maximum likelihood. As logarithms are taken when calculating the negative binomial likelihoods under a multiplicative specification, we add one to all patent numbers before estimation.

The number of lags was determined by estimating the equation with up to twenty four lags. The Akaike Information Criteria and the coefficient significance of the highest lagged term were used to determine the optimal number of lags. The results are shown in the appendix. Four lags were taken as optimal.

If shocks to the patent process are persistent, the accumulating error may be collinear with a change in the time trend and estimator distributions may be different from their non-persistent versions (see Hamilton (1994, pages 497-501) for derivations with normally distributed errors). Persistent shocks may be misidentified as a break in the series. We tested for persistency in expected patent rates using a Zivot-Andrews test on the logarithmic series, which tests for unit roots when there may be a trend change or unit shift in the series. The statistic was -5.42, which is significant at five percent against a null hypothesis of a unit root.

We estimate the equation repeatedly with the value of u varying across the period under consideration. Estimations are not made for the first and last six months. Estimation is undertaken for the equation with a unit shift alone, a trend change alone, and both break types.

The procedure produces a series of p-values associated with an asymptotically valid normal test that the break variables are individually or dually equal to zero at each date. The p-values over the eighteen month period from July 2002 to December 2003 are directly compared to find the minimal one, and the parameter estimates at the corresponding break date calculated. The validity of the direct comparison relies on the assumption that our prior information is that a patenting incentive shock is equally likely to occur at any point in the period, implying a rising probability of incentive shock conditional on no shocks prior to the considered date.

The p-values may not accurately reflect the probability of a break in the generating process because of parameter biases in the finite sample maximum likelihood estimations. Moreover, as estimates of the probability that the patent process has a significant break during the eighteen month period, they do not allow for our search procedure. We want the probability of no break at the date conditional on that date having the greatest unconditional evidence of a break in the period, rather than the unconditional probability of no break at the date. To correct for MLE bias and allow for conditionality, we estimate the patent process without breaks and use the estimated parameters to generate a thousand Monte Carlo simulations of the process. For every simulation, we test for each type of break over the period and record the minimal p-values. The frequency of the simulated p-values lying below the actual data p-value is taken as an adjusted p-value for the break.

Our second empirical specification considers multiple shocks over the period, with possibly repeated entry, exit, and return to a fixed number of patenting incentive states. The incentive states have expected patent numbers with the autoregressive coefficients constrained to zero, that is,

$$E(P_t|I_t, R_t, H_t) = A_{j(t)}e^{b_{j(t)}t}$$

where $A_{j(t)}$ and $b_{j(t)}$ are the coefficients in state $j(t)$, the state prevailing at time t. The coefficients are specific to each state and do not have any relation to each other across states. The stochastic distribution of the patents is Poisson in each state. When in one state, the probabilities of moving to the same or another state in the following time period depend only on the current state, not the transition history or the date. These assumptions define a Poisson hidden Markov model (Zucchini and MacDonald, 2009).

We estimate the parameters in the second specification by maximum likelihood estimation. The most likely sequence of states is also estimated by maximum likelihood using the Viterbi algorithm. The algorithm calculates the most likely sequences of increasing lengths, given the estimation transition matrix and observed data, and ending in each possible state, using the Markov property to increment the length. For the sequences extending over the whole period, the maximum likelihood sequence is found by inspection of the likelihood of the sequences for each end state. Zucchini and MacDonald (2009) describe the mathematics of the algorithm and give a computer code implementation.

We estimate the model for two and three states. The first state in any estimation is liable to capture the transition stage at the USPTO between obligatory and optional secrecy in patent applications, when reported patents are limited. It may also be applicable to describe later low levels of patenting, so it can serve a purpose in describing patenting after the transition process has ended. We also estimated a four state model, which potentially would be suitable to describe the transition stage, then a pre-SOX patenting rate, then an immediate post-SOX rate, and finally a long term stage. However, on our main data the likelihood maximising four state HMM used only three states, and so the estimates are included within our three state results.

For a sequence of n periods, the number of state sequences is 2^{n} or 3^{n}. The most likely path differs little from many other paths, and its individual probability may be low. The estimates may be influenced by convergence thresholds even if they are small. There are plausibly multiple local maxima that may be identified by maximum likelihood. In order to ensure that the estimated sequence reflects a global maxima and its broad shape has a high concentration of probability among potential sequences, we repeat the estimation for multiple starting values for the Poisson rates in each state, selected either by visual inspection of plausible values or as starting, average, and end values of second or third quantiles of the data. The results we present are representative of the shape and timing of the shifts between states across most estimations.

The estimation was performed in the R computer language (R Development Core Team, 2009) using the library packages Hmisc, urca, and MASS for the p-value estimation and testing, and the msm package for the HMM estimation. All R code for the modelling is available at the author's website[1], together with base data and simulated p-values.

5 Results

5.1 US residents' patent applications

Figure 2: P-values for breaks in US residents' applications by date. The solid line shows values for dual breaks, the dashed line for unit shifts, and the dotted line for trend changes.

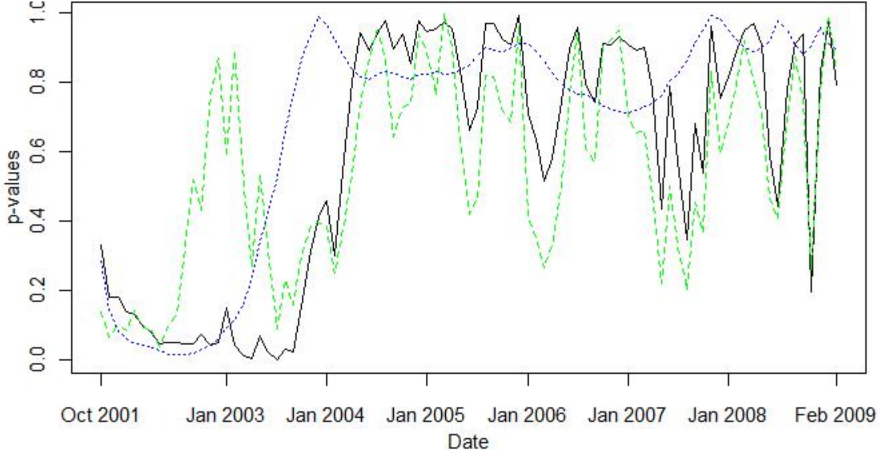


Figure 2 shows the p-values for unit, trend, and dual breaks (that is, breaks in either the unit or trend component) at dates throughout the period 2001 to 2009 for published patent applications by US residents. The solid line indicates values for dual breaks, the dashed line for unit shifts, and the dotted line for trend changes. The unit shifts are most likely to have occurred around the start of 2002, followed by the middle of 2003. The trend changes are most likely to have occurred in the middle of 2002. Changes elsewhere are unlikely. After allowing for correlations between the breaks, the dual break curve finds strongest evidence for a break occurring at dates from around the middle of 2002 to the last half of 2003. The strongest evidence occurs around the second quarter of 2003.

Table 3 shows the coefficient estimates for each break date. The coefficients are evaluated at the minimal p-value from July 2002 to December 2003. The unit shift is most significant at July 2003, where it has a negative coefficient. It is significant at ten percent,

---

[1] http://ebasic.easily.co.uk/02E044/05304E/sox_effect_on_innovation.html

Table 3: Coefficient estimates for breaks at the time of the minimum p-value in July 2002 to December 2003 inclusive

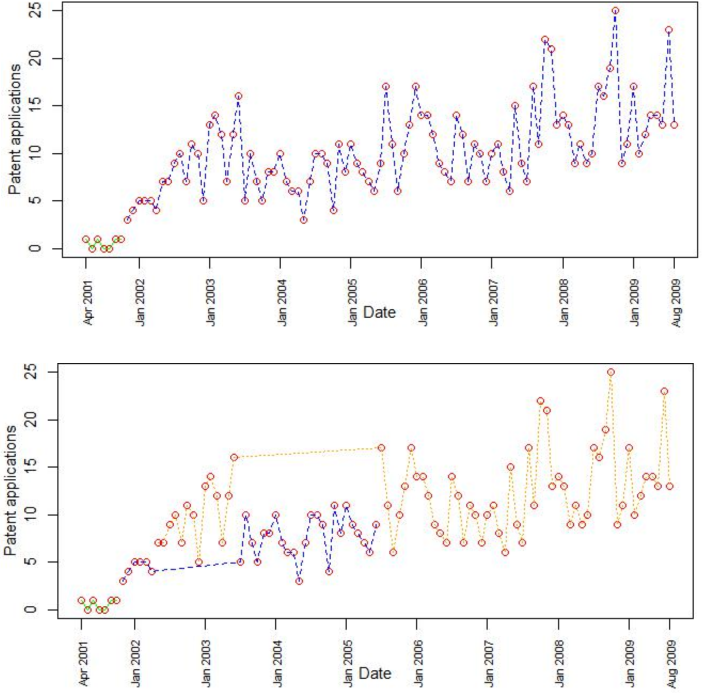
	1	2	3
Unit shift	-0.198 *		-0.643 ***
	0.116		0.176
Trend change		-0.115 **	-0.059 ***
		0.047	0.018
Time	0.006 ***	0.12 **	0.068 ***
	0.002	0.048	0.019
Constant	1.125 ***	0.038	0.74 ***
	0.205	0.497	0.245
Dependent(t-1)	0.285 ***	0.238 **	0.208 **
	0.099	0.101	0.101
Dependent(t-2)	0.008	-0.067	-0.069
	0.1	0.104	0.102
Dependent(t-3)	0.269 ***	0.186 *	0.156
	0.099	0.104	0.104
Dependent(t-4)	-0.095	-0.209 **	-0.218 **
	0.092	0.099	0.099
Break significance	0.087	0.015	0.001
Adjusted significance	0.522	0.084	0.017
Pseudo R^2	0.52	0.54	0.58
Box-Ljung p-value	0.98	0.9	0.89
Break date	Jul-03	Jul-02	Jul-03

Standard deviations are shown below the coefficients. *** denotes an unconditional p-value of less than 0.01, ** of less than 0.05, * of less than 0.1. Time is measured in number of months elapsed since March 2001.

but the adjusted p-value shows the coefficient has no significance. The minimal trend p-value is at July 2002, where the change is negative. The change is equivalent to an eleven percent reduction in expected patents compounded over time. It has an adjusted significance of eight percent. With both breaks included, the minimal p-value occurs in July 2003. The unit shift coefficient has a negative sign, marking a one-off reduction of 47 percent in the expected rate of patenting, while the trend change is also negative, equivalent to a six percent compounded reduction in the patenting rate. The adjusted probability that the coefficients are simultaneously zero is two percent.

The top panel of figure 3 shows the decoded states of a two state hidden Markov model applied to the patent data. There are two temporally connected states identified. The first runs for seven months from April to October 2001 and has low and roughly constant

Figure 3: Decoded states for a two state hidden Markov model (top) and three state HMM (bottom) for US residents' published patent applications; states are connected by coloured lines.



patenting rates. The second runs unbroken from November 2001 onwards and has an apparently increasing patent rate. The bottom panel of figure 3 shows the decoded states for a three state HMM process. The early connected state to October 2001 is unchanged. A second state from November 2001 persists until April 2002, and then returns for a connected two year period from July 2003 to June 2005. The second state's sequence is interrupted by a third state which lasts unbroken until the second state is restored in July 2003. After exit from the second state in June 2005, the third state returns and persists until the end of the period. The third state has a higher apparent patenting rate than the second state after transition. It is not clear whether the time trend is higher or lower.

Table 4 contains the maximum likelihood estimates of the coefficients in each Poisson patenting state, showing the constant and time trend components of the expectation. The short-lived first state has a low rate throughout its duration under both the two state and three state specifications. The second state has a far higher constant component to the state expectation. For the three state estimation, the last state has a higher constant component but lower trend than the second state.

The observations indicate a large upwards shift in the patenting rate and trend in

Table 4: Coefficient estimates for the states in the two and three state HMMs on US residents' applications

	1	2
State 1		
Constant	0.209	1.004
Trend	-0.234	0.141
State 2		
Constant	9.819 ***	8.616 ***
Trend	0.088 ***	0.135 *
State 3		
Constant		11.194 ***
Trend		0.062
Date of first entry into state 1	Apr-01	Apr-01
Date of first entry into state 2	Nov-01	Nov-01
Date of first entry into state 3		May-02

*** denotes an unconditional p-value of less than 0.01, ** of less than 0.05, * of less than 0.1. The p-values are for an asymptotically valid chi-squared test using parametrically bootstrapped standard deviations with 100 bootstraps. Time is measured in number of months elapsed since March 2001 multiplied by 0.1.

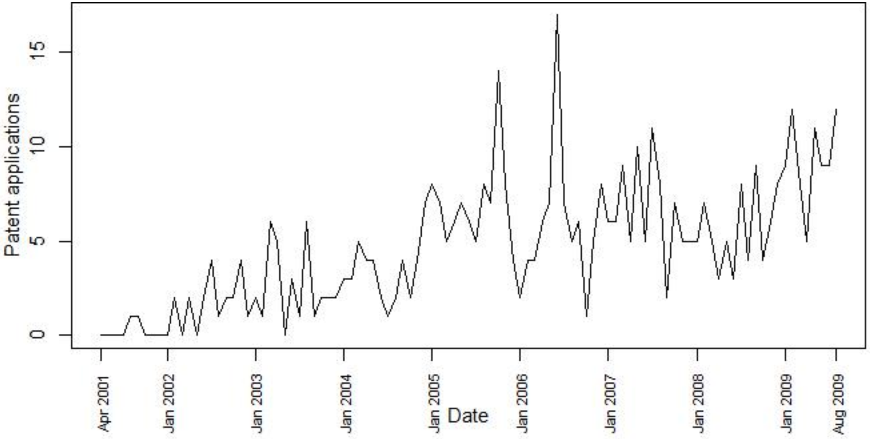
November 2001 away from the negligible patenting activity that existed until then. A further upwards shift in level, if not in trend, occurred in May 2002. The patenting state that emerged was interrupted temporarily from July 2003 to June 2005 by the restoration of the previous state. The timing of the July 2003 break is consistent with an adverse effect due to SOX. We thus find evidence in support of the patenting reduction after SOX described in hypothesis one from both the break p-values and from the three state HMM model. There is additionally evidence from the HMM model of a patenting correction at a timescale consistent with hypothesis one, and that SOX has no persistent net effect.

5.2 US and non-US resident patenters

A number of explanations arise from factors affecting the patenting of both US and non-US residents. Among them are an exhaustion of commercialisable ideas in stem cell technologies, a shift of global investor preferences against funding its innovation, a shift against speculative investment generally, a reduction in US demand for stem cell technologies, processing delays at the US patent office, or an increasing proportion of applicants choosing to publish their applications over time.

Under this class of explanations, the incentives or ability of non-US residents to patent

Figure 4: Published patent applications by month for non-US residents



in stem cell technologies is reduced at the same time as those of US residents. A break in non-US residents' published patent applications is likely to occur and approximately coincide with the previously observed breaks. The earlier empirical equations form the basis for testable hypotheses on a break in their published patent applications.

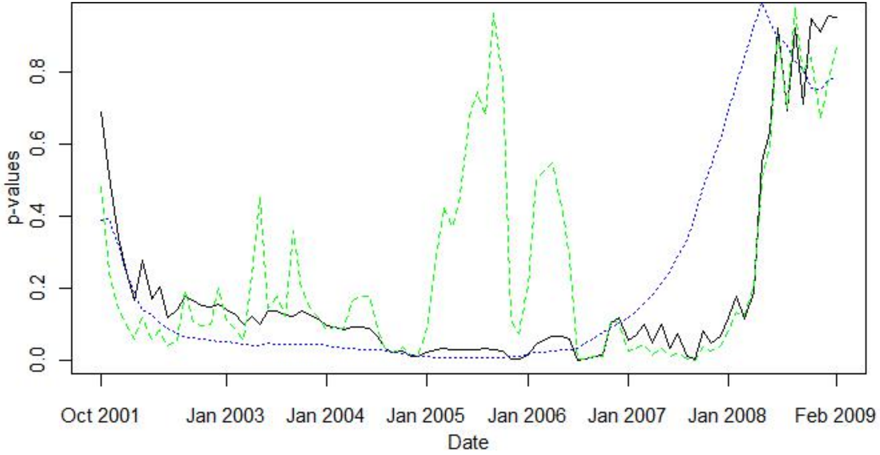
The data for non-US residents is also from the US Patent and Trademark Office. It is shown in figure 4. The publication rate rises until around the start of 2003, then stabilises until the end of 2004 when it begins to increase rapidly, lasting until the middle of 2006. It undergoes a possible downwards shift in the level of the series. It then rises, falls, and then rises again until the middle of 2009. The observed pattern is not clear; it could alternatively be explained as a roughly constant growth rate with a period of greater volatility around the end of 2005 and start of 2006.

Figure 5 shows the p-values for unit shifts (dashed line), trend changes (dotted line), and dual breaks (solid line), estimated with an autoregressive lag of two. Unit shifts are most likely at the end of 2004, and in the second halves of 2006 and 2007. The period from the middle of 2002 to the end of 2003 has relatively less evidence of a break. Trend changes are most likely to have occurred in 2005. The most probable times for a break in at least one of the series are at the ends of 2004 and 2005, and the middle of 2006. The SOX period is relatively less likely to have experienced a break.

The results contrast with the observed decline in the US residents' patent rate at the time of SOX. The two results are consistent with the decline's cause acting only on US residents. They provide evidence against the class of alternative explanations to SOX that explain the decline in terms of globally effective factors.

The decoded states for a two state hidden Markov model of the non-US patent data are in the top panel of figure 6. The initial state lasts from April 2001 to January 2002. From

Figure 5: P-values for breaks in non-US residents' applications by date. The solid line shows values for dual breaks, the dashed line for unit shifts, and the dotted line for trend changes.



February 2002 it undergoes an extended break and is restored only in September 2007. It then endures to the end of the period. The intervening second state lasts unbroken for over five and a half years. It has an apparently higher level than the first state. They both have an upward trend. In the three state model shown in the bottom panel, the first state is unchanged from the two state model. The second state only occurs for the period from February 2002 to November 2004. It has a higher apparent rate than the first state. After its end until the restoration of the first state, a third state occurs with apparently even higher rate and uncertain trend.

We do not see any evidence of a downward break in the period around SOX, or any break at all. The analysis produces a similar conclusion to the p-values break results. It is less likely that globally applicable factors were the main cause for the reduction in US residents' patent rates, and more likely that SOX was.

5.3 Institutional and non-institutional patenters

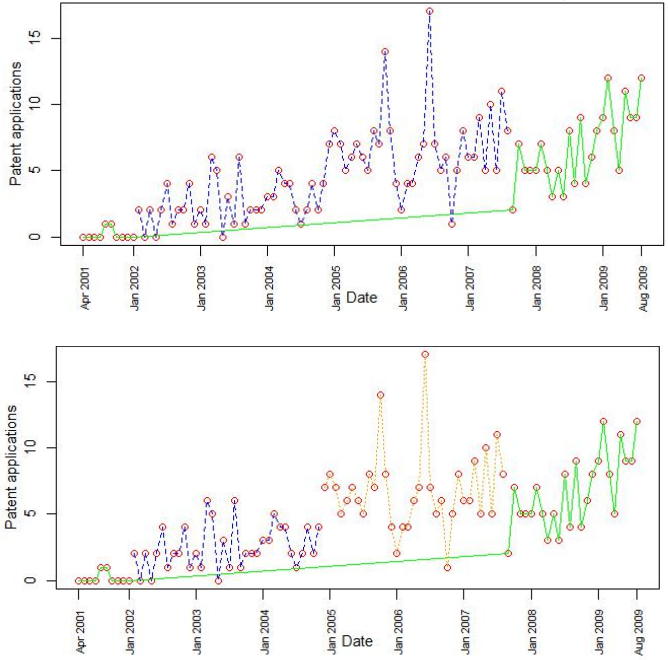
SOX only applies to publicly traded companies. If it was responsible for the break and subsequent correction found in the US resident patent series, and if patent sources other than public companies were insulated from its impact, we would not expect to see similar behaviour in their series. In this section, we examine patenting by institutions including government and universities. [2]

US patents record the name of the inventor or inventors and the legal assignee of the patent if they differ. When separately specified in our data, the legal assignee is generally

---

[2] Thanks to Pelin Demirel for suggesting comparison of institutional patenting and exclusion of technology specific behaviour.

Figure 6: Decoded states for a two state hidden Markov model (top) and three state HMM (bottom) for non-US residents' published patent applications; states are connected by coloured lines.



an institution or company. We record the assignee for all US residents' patent applications over 1986-92. Of the 982 applications, 408 include an assignee. There are 237 assignees with the highest number of patents due to a single assignee being 22. The assignees are then separated into two groups of institutional and non-institutional patenters. The criteria for inclusion in the former group is that they are a university, the US Government, a hospital or group of hospitals, a medical centre, a foundation, or an institute. There is also a single patent assigned to a person that we include in the institutional group on the grounds that they too would not be directly affected by SOX. The other named assignees are all corporate (both private and public) and are put into the non-institutional group. There are 211 institutional patents and 197 non-institutional patents.

Some of the organisations in the institutional patenter groups operated through incorporated vehicles, and may potentially have been subject to SOX if they were public companies. We assume that this possibility did not apply to many of the organisations and so did not distort the exposure. The assumption is reasonable as the institutional patents were largely from universities (155) rather than non-universities (56).

After data processing, we estimated the state sequences in three state hidden Markov models on the institutional and non-institutional patent series. The results are shown in

Figure 7: Decoded states for a three state hidden Markov model for US residents' published patent applications from institutions (top) and non-institutions (bottom); states are connected by coloured lines.

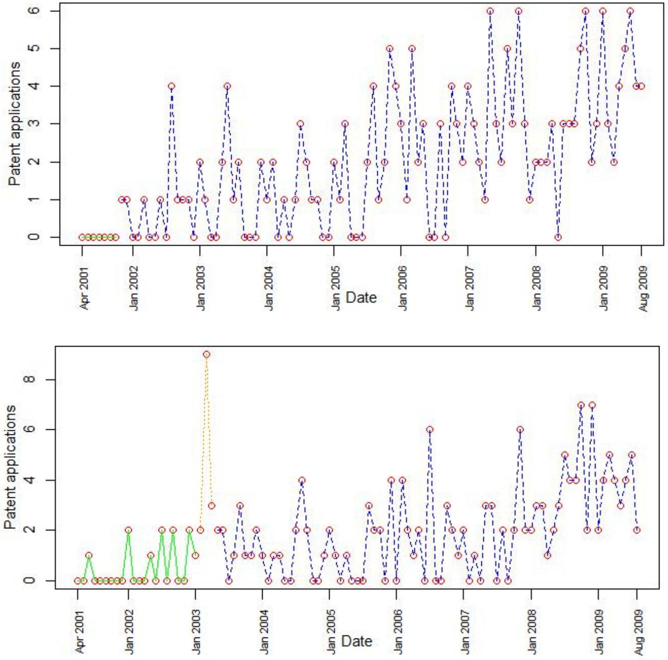


figure 7. As previously, patent numbers are collected into states and connected by coloured lines.

The estimated states for institutional patenters are shown in the top panel. Only two states are identified. The first consists of a seven month period from April to October 2001 when there were no patents and the reported patent applications are affected by the change in patent secrecy described in section 3. From the end of 2001 onwards a single state is identified to describe the general trend upwards in patenting. No third state required, and so there is no change in the generating state around the time of SOX. Visually, there is possibly a slight drop in patenting in the second half of 2003, but state sequences with a state change are less likely than the single state sequence determined. It is possible that the slight drop may be due to a SOX spillover from public companies, for example if institutions patent anticipating licensing or intellectual property sales to public companies.

In the lower panel, we see the estimated states for non-institutional patenters. There is an initial temporally connected state of low patenting that runs to January 2003. Then there a short-lasting state for three months which contains the highest spike in patent numbers over the whole 2001-9 period. Then follows the final state that is initially lower than the previous state but rises over time. While the transition times are consistent with a SOX effect, the state switching differs from the $3-2-3$ sequence observed for total US originating patents. Moreover, the existence of the second state is determined by the patent spike in March 2003. When we set the patent value in the month to zero, we obtain almost the same state sequence as for institutional patents with two states transitioning once in January 2002.

Caution is due in interpreting the state sequences because the number of patents in any month is small, particularly early in the period. Caution taken, we can still note it is the presence of patents without assignees that determines the behaviour of total US originating patents, with the post-SOX fall and rise. It is possible that the non-assigned patents are speculative work that organisations find it efficient to allow their workers to pursue independently, and after SOX major companies are taking less risk and concentrating on core work. We investigate further in section 5.5.

5.4 Exhaustion of ideas in the US

A further explanation for the observed break is based on the pattern of emergence and exhaustion of patentable ideas. Ideas are posited to lead to patents. Ideas start slowly at first as a technological idea is first investigated, then as the most promising prospects are discovered the generation of patentable ideas accelerates, and finally the prospects begin to be exhausted and the rate flattens. The observed break corresponds to the beginning of the gradual exhaustion and deceleration.

In order to distinguish this explanation from the general global explanations just discussed, we have to posit that the decline is in the US alone. The situation might arise if US innovators are innovating in a different type of patents than foreign innovators, for example if they are far more technologically advanced. The condition requires that technological innovation does not rapidly spill across borders, which is open to theoretical and empirical question.

The explanation suggests a trend change in the patenting rate without a level shift in it. The time-varying p-value analysis earlier found evidence for a trend change but no significant evidence of a unit shift. The hidden Markov model estimation by comparison found that there was a temporary downwards level shift in the aftermath of SOX. Thus, the evidence available so far provides mixed evidence against the exhaustion hypothesis.

We investigate further by examining behaviour after the downturn. Under the exhaustion hypothesis, the rate of innovation will decline further after the initial transition. We test empirically for the presence of further downturn by modifying the form of expected patents to

$$E(P_t|I_t, R_t, H_t) = \exp(a + bt + et^2 + d_2 DT_{t>u}) \prod_{i=1}^{\infty} P_{t-i}^{c_i}$$

where the period assessed is from August 2003 to August 2009. Our first specification has the restriction $e = 0$ with u set at the mid point between the break date and the end of the period, ie August 2006 (so allowing for a late break). Our second specification has $d_2 = 0$ (allowing for a quadratic downturn in time with linear time as a covariate), and our third has $b = 0$ and $d_2 = 0$ (allowing for a quadratic downturn in time without a linear covariate).

Table 5 shows the sign and significance of the late trend change and quadratic terms in the maximum likelihood estimations. The sign on the late trend change is negative, but the coefficient has low statistical significance. The same is true for the quadratic time term in the presence of a linear time trend. If quadratic time is present without a linear time trend, it has a positive coefficient and is highly significant, probably capturing some of the linear trend effect. In summary, we again find limited evidence for the exhaustion hypothesis as it is presented here.

Table 5: Coefficient estimates for the states in the two and three state HMMs on US residents' applications

Variation from a linear process	Sign, p-value of the non-linear term
Late trend change	-, 0.63
Month^2 (with covariate month)	-, 0.55
Month^2 only	+, 0.00

5.5 The earlier market downturn and reduced patenting incentives

The observed change in the patenting process may alternatively be attributed to a downturn in market prospects. Under this hypothesis, there was a loss of confidence or funding in the market, and many innovators exited the market. There is data presented in Lysaght ***et al.*** (2008) showing a decline in the financial performance and size of the tissue engineering market prior to SOX, although the stem cell sub-sector experienced employment growth and was financially sheltered by its pre-commercial status.

After a market decline and financial rationing, if remaining financiers are reasonably accurate at assessing financial prospects of innovations the residual funding should be allocated to innovations offering better financial prospects than innovations funded prior to the rationing. By contrast, SOX has been associated with reduction in corporate funding for new, small, and high growth companies, who may be expected to produce innovations with better than average, if risky, prospects. Thus, the SOX and market downturn explanations for patent number changes have different implications for the prospects of the remaining innovations relative to innovations prior to the downturn.

We may noisily measure the financial prospects of patents by the number of patents citing them. The propositions and measure allow us to form a testable hypothesis associated with the downturn conjecture, that the rate of citing of new patents was higher after the process change than before it. The alternative hypothesis associated with a SOX effect is that the rate of citing was the same or lower after the process change.

Table 6: Rate of citing of patents whose applications were filed April 2001 to March 2002 or April 2003 to March 2004 inclusive, split by issue year

	Applications filed in 2001/2		Applications filed in 2003/4	
Issue year	Counts	Average citations	Counts	Average citations
2002	1	1		
2003	5	4.8		
2004	15	5.3		
2005	15	2.5	1	0
2006	17	1.3	5	0.6
2007	6	2.2	5	1.2
2008	3	1	12	0.2
2009	5	0	9	0.1
2010	1	0	8	0
Total	68	2.6	40	0.3

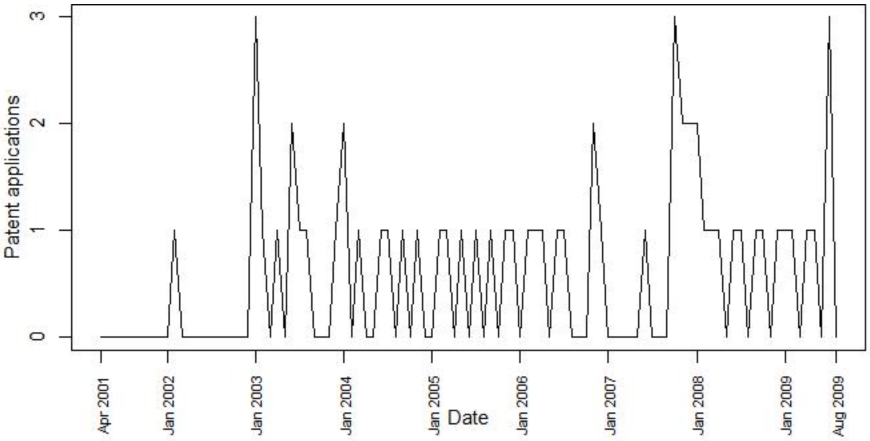
Data is from the US Patent and Trademark Office. The website search term was, for example, TTL/("stem cell" OR "stem cells") and APD/4/1/2001->3/31/2002 and ISD/4/1/2006->3/31/2007.

Table 6 shows the citation rates for stem cell technology patents whose applications were filed from April 2001 to March 2002 inclusive, compared with those filed between April 2003 and March 2004. Data is from USPTO (2013). The rates are given for all citing patents up to July 2010, and are split by the year of patent issue and for all issue years combined. The combined citation rate is far higher for the earlier period of filing. The higher combined rate may be possibly attributable to a longer post-issue period for earlier filed patents and hence more exposure and opportunities for citation. The length of the period since issue is controlled by the comparison of patents issued in the same year. These rates are consistently higher for the patents filed in 2001 to 2002. We find no evidence of an increase in the citation rate for patents filed after the observed break, or that financial prospects as proxied by citation rates picked up after it. The result weakens support for the hypothesis that a financial downturn was the main driver of the observed changes in the patent process.

5.6 US Government funding cuts for human embryonic stem cell research

A further candidate explanation for the observed change in the patenting process is that legislative or federal funding activity relating to human embryonic stem cell technologies in the United States affected the rate of patent applications. In 2001, the US administration restricted federal funding for research in human embryonic stem cells (United States White House, 2001). The decision may have driven down innovation in human embryonic stem cell technologies directly, and indirectly affected companies producing other stem cell technologies (see Salter (2005, page 8) for a description of a contamination effect between the two forms of technology).

Figure 8: Published patent applications in human embryonic stem cells technologies by month for US residents



The mechanism is feasible, but it seems implausible that it would have a sufficiently large effect to be the only cause of the observed process change. Firstly, federal research funding was restricted on stem cell lines from embryos destroyed after the date of the restriction, but not prior to the date. Thus, the restriction was partial. Secondly, it did not affect privately funded research. Thirdly, the number of human embryonic stem cell patents was very small compared with the number of general stem cell patents. From a search on the US Patent and Trademark Office database, the cumulative number of US residents' patent applications between April 2001 and August 2003 with "human embryonic stem cell" or "human embryonic stem cells" in the title was just ten compared with 185 general stem cell technology patents. Even if the federal funding restriction stopped most innovation in human embryonic stem cell technology, the direct effect on the broader innovation group would be very limited. Fourthly, if there is a contamination effect on non-embryonic innovation from the restriction, it would have to be very large and persistent to account for the observed process change. There seems little prior reason to believe in such an effect. Fifthly and finally, insofar as it is possible to determine a process shift from the low numbers of human embryonic patents in figure 8, it does not seem to be any greater than that affecting the general stem cell patent series.

5.7 Breaks in other patent series

If SOX explains the changes in the US stem cell patenting series, then similar changes should be observed in the series of other new and risky technologies. We briefly test for them in this section. New data is extracted from the US Patent and Trademark Office online database, modifying the search criteria to

TTL/("product" OR "products") and PD/4/1/2001->4/31/2001 and ICN/("US").

where "product" is a product or product characteristic. We select a handful of terms from SIC codes to give a variety of technology characteristics. Thus, "product" is set to be "furniture", "telecommunication", "computer", and "medical". Considering the length of time and stability of the products' respective markets, the size of innovative companies in them, and the chance of product failure, we may guess that the list is in increasing order of response to SOX.

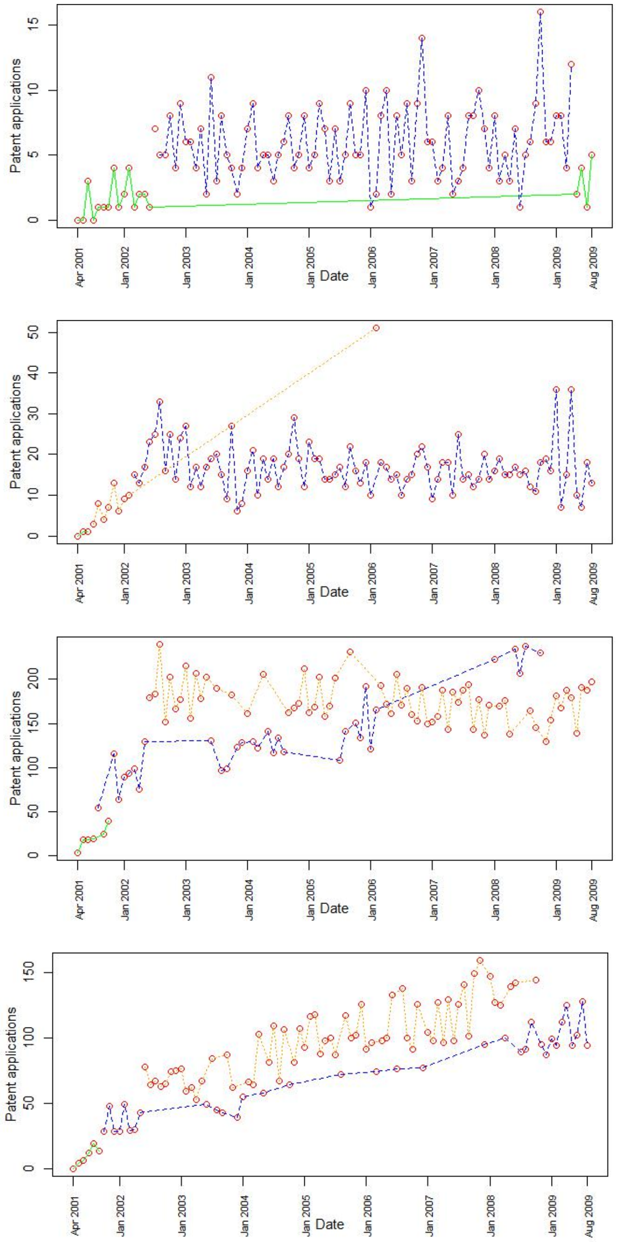
Figure 9 shows the results of running a three state HMM on each of the resulting product series. The top panel shows the estimated states for furniture patents. There is an initial period of low patenting, when the secrecy policy of the USPTO was adjusting. There is subsequently a higher single state from the middle of 2002 until April 2009, when the low initial state returns.

Estimated states for telecommunication patents are in the second panel. There is an initial state characterised by low but quickly increasing patents. Then a single state continues until August 2009, with the exception of a single month in 2006 which has a patenting rate much higher than in other months and so returns to the initial state with its patents extrapolated. Visual inspection indicates a possible downward trend in patents from the middle of 2002 to the end of 2003 with a recovery over the following year, but maximum likelihood doesn't indicate that it should be represented through a different state.

The third panel displays the estimated states for computer patents. There is an early state with low and rapidly rising patent numbers. From the middle of 2001 a higher level but more slowly growing state emerges, which is broken in June 2002 by a third state at an even higher level. The state persists until the summer of 2003, when the second lower state returns and is dominant until the last part of 2004 and for much of the period after 2006. The second state returns in 2005 and 2008. The interpretation of the state fluctuations is not entirely clear, but the 2003-4 reduction and restoration is consistent with an effect due to SOX.

In the bottom panel we see estimated states for medical patents. After a transient high growth, low level state, a second state emerges that occurs uniquely from the final quarter of 2001 to the middle of 2002. Then a higher third state emerges which persists until May 2003, when it begins a nine month oscillation with the lower state. From the middle of 2004, the higher state is dominant with occasional re-entries into the second state. The second state returns to be the principal state from the middle of 2008.

Figure 9: Decoded states for a three state hidden Markov model for US residents' published patent applications mentioning furniture (top), telecommunication (second), computer (third), and medical (bottom); states are connected by coloured lines.



The behavioural response across all products has been broadly as expected. Depending on the riskiness of the market, patenting falls at a time consistent with an effect due to SOX, and is restored subsequently as is consistent with a decline in uncertainty about it. The effect is absent for the lowest risk market and strongest for the highest risk market.

The results are as we expected, but strong cautions should be made. The selection of products to examine is ad hoc, though motivated by diversity and without exclusion of trialled product analyses. The classification of market risk is also informal and could be challenged. For these reasons, this section should be considered an outrider for a more formal and exhaustive selection, processing, and analysis.

6 Conclusion

We have presented and solved for the investment choice faced by investors in an innovative company before SOX, immediately after it, and later when market uncertainty has reduced. Our hypothesis was that a temporary downwards shift in innovation occurred after SOX, and analysis of stem cell technology patents found supporting evidence. Our theoretical analysis further suggested that SOX's long-term effect depended on the relative sizes of SOX costs and the risk premium associated with uncertainty about managerial quality. Empirical analysis indicated that SOX had no long term effects on innovation in stem cell technology. Thus, we agree with the literature emphasising transient uncertainty effects of SOX (Akhigbe ***et al.***, 2008) and further find its long term neutrality.

The literature has indicated that managerial incentives changed after SOX (Cohen ***et al.***, 2013). We did not model managerial incentives in our theoretical presentation. To an extent, they could be captured within our model as managerial perceptions of uncertainty due to SOX would also diminish over time. However, there would be a misalignment between managerial and investor incentives and risks as managers would be able to observe their own management quality prior to SOX and investors would not, and changes in incentive alignment due to SOX could modelled more fully. For a full welfare analysis, investor, manager, and consumer utility would have to be considered.

Our theoretical presentation suggested how innovation delays respond to SOX's characteristics. An optimal form of implementation could be calculated based on trade-offs between speed of implementation, expected costs, and dispersion of costs. An implementation model could be estimated or calibrated to SOX data, and counterfactuals could be investigated. A detailed dynamic programming model could give a more complete method of exploration. Other disclosure regulations, that like SOX generate real options for investors, could be used to get greater data breadth in estimation.

Various small departures from the empirical model's assumptions were tested during this paper's preparation, being rejected as imperfect models of count data although they may be more valid for other data sets. We tested additive instead of multiplicative patent forms and used random walk processes for patent emergence with sandwich and GARCH errors. The conclusions were not altered.

Most analyses of SOX's effects are based on panel data rather than a single time series. We may extend the approach here to patent panel data, using multivariate count data and HMM models which are established in the literature. The dummies that are used in the SOX literature based on differences-in-differences may not enter these count data models in additive form making their removal more complex, if possible at all. Thus, relative to the existing literature it may be more difficult to isolate SOX breaks and separate SOX effects from other effects, and the graphs presented here may not readily translate into a panel data analysis for visual examination. There would also be the problem that series breaks could occur at different times. Again, a dynamic programming model could be conceivably used, or a cellular automata model.

We used publication date rather than application date when constructing the patent time series, in order to ensure that late period values were not distorted by delayed publication of applications. As more time passes from the end of the period, we can have more confidence that such distortions will not occur, so future studies could use application dates. Preliminary examination of the US stem cell patent series with application dates indicates that a temporary drop was observed after SOX, in line with the results above, but it is less clear than with publication dates. A multi-series analysis may be required to obtain statistical significance in changes.

Other future empirical work could examine SOX-related causes for the observed patenting shifts in more detail, in order to determine whether the mechanism suggested here or an alternative was responsible. For example, one alternative would be companies avoiding patenting in order to avoid value creation and so remain below the threshold for accelerated SOX filing (see Gao ***et al.***, 2009).

Our work suggests some policy implications, based on the restricted model we have presented. Other things being equal, unnecessary uncertainty about the implementation costs of a reform increases investment delays so should be avoided, for example by keeping investors informed about the details of legislation, particularly in newer, smaller, and more volatile industries. Unnecessary uncertainty may also be avoided by following a timely legislative process. Holding other things equal, increased implementation costs increase delays and lower long term innovation, while disclosure legislation increases long term innovation only if its costs are below the investor's risk premium associated with ignorance about the non-disclosed information. We have neglected improvements in managerial behaviour that may modify this conclusion and which could be examined separately or connected with our study in future work. A more rapid implementation may increase investment delays during the implementation period. These characteristics are unlikely to operate independently, for example as faster implementation may be likely to bring increased costs and uncertainty about the costs, so that a fuller model looking at trade-offs and counterfactual implementations would be valuable in policy discussions.

On a final policy note, our finding of no net long term effect of SOX on innovation does not imply that Sarbanes Oxley had no impact. Our theoretical model suggested that the net effect on innovative investment can be decomposed into SOX's effects on costs and on investor risk premiums linked to uncertainty about managerial quality. SOX may have individually impacted on both of them so that the effects were equal and opposite. It is therefore feasible that either costs or investor risk premiums could be reduced in order to increase the net benefits of SOX, through measures such as the Jumpstart Our Business Startups Act of 2012 (Biotechnology Industry Organization, 2012).

Appendix: Autoregressive diagnostics

Table A1: AIC and highest lag p-values by AR order for negative binomial models of patent applications

	US applicants		Non-US applicants	
Lag	AIC	Lag p-value	AIC	Lag p-value
1	424.6	0.09	375.1	0.05
2	426.6	0.9	374.4	0.1
3	427.5	0.31	376.2	0.61
4	426.7	0.08	378.1	0.83
5	428.6	0.77	379.9	0.67
6	430.3	0.62	381.9	0.98
7	432.1	0.66	382.6	0.24
8	432.6	0.21	382.2	0.13
9	434.3	0.63	383.6	0.44
10	436.3	0.95	384.5	0.28
11	437.4	0.33	386.4	0.73
12	438.9	0.49	387	0.24
13	436.6	0.04	388.3	0.42
14	434.5	0.05	390.3	0.93
15	435.8	0.38	384	0
16	434.5	0.07	386	0.96
17	436.5	0.94	386.5	0.23
18	434.7	0.05	386.8	0.19
19	435.7	0.32	388.5	0.62
20	437.4	0.58	389.2	0.25
21	438.5	0.34	391.1	0.76
22	437.9	0.1	393.1	0.93
23	438	0.17	394.8	0.57
24	433.3	0.13	390.8	0.17

References

Ahmed, A., McAnally, M., Rasmussen, S. and Weaver, C. (2010) How costly is the Sarbanes Oxley Act? Evidence on the effects of the Act on corporate profitability, *Journal of Corporate Finance*, **16**, 352–69.

Akhigbe, A. and Martin, A. (2008) Influence of disclosure and governance on risk of US financial services firms following Sarbanes-Oxley, *Journal of Banking & Finance*, **32**, 2124–35.

Akhigbe, A., Martin, A. and Newman, M. (2008) Risk shifts following Sarbanes-Oxley: Influences of Disclosure and Governance, *The Financial Review*, **43**, 383–401.

Akhigbe, A., Martin, A. and Newman, M. (2010) Information Asymmetry Determinants of Sarbanes-Oxley Wealth Effects, *Financial Management*, pp. 1253–72.

Archibugi, D., Filippetti, A. and Frenz, M. (2013) Economic crisis and innovation: Is destruction prevailing over accumulation?, *Research Policy*, **42**, 302–14.

Ashbaugh-Skaife, H., Collins, D., Kinney, W. and LaFond, R. (2009) The Effect of SOX Internal Control Deficiencies on Firm Risk and Cost of Equity, *Journal of Accounting Research*, **47**, 1–43.

Bargeron, L., Lehn, K. and Zutter, C. (2010) Sarbanes-Oxley and Corporate Risk-Taking, *Journal of Accounting and Economics*, **49**, 34–52.

Biotechnology Industry Organization (2006) BIO Issues Statement On Anticipated Changes to Sarbanes-Oxley Implementation, Downloaded from http://www.bio.org/media/press-release/bio-issues-statement-anticipated-changes-sarbanes-oxley-implementation.

Biotechnology Industry Organization (2007a) Biotech Research and Development Will Benefit from Sarbanes-Oxley Changes, Downloaded from http://www.bio.org/media/press-release/biotech-research-and-development-will-benefit-sarbanes-oxley-changes.

Biotechnology Industry Organization (2007b) Greenwood to Congress: Proposed Sarbanes-Oxley Reforms, Downloaded from http://www.bio.org/media/press-release/greenwood-congress-proposed-sarbanes-oxley-reforms.

Biotechnology Industry Organization (2012) BIO Lauds Signing of JOBS Act into Law, Downloaded from http://www.bio.org/media/press-release/bio-lauds-signing-jobs-act-law.

Blundell, R., Griffith, R. and Van Reenen, J. (1995) Dynamic Count Data Models of Technological Innovation, *Economic Journal*, **105**, 333–44.

Cheng, M., Dhaliwal, D. and Zhang, Y. (2013) Does investment efficiency improve after the disclosure of material weaknesses in internal control over financial reporting?, *Journal of Accounting and Economics*, **56**, 1–18.

Chhaochharia, V. and Grinstein, Y. (2007) Corporate Governance and Firm Value: The Impact of the 2002 Governance Rules, *Journal of Finance*, **62**, 1789–825.

Cohen, D., Dey, A. and Lys, T. (2013) Corporate Governance Reform and Executive Incentives: Implications for Investments and Risk-Taking, *Contemporary Accounting Research*, accepted article.

Doidge, C., Karolyi, A. and Stulz, R. (2009) Has New York become less competitive in global markets? Evaluating foreign listing choices over time, *Journal of Financial Economics*, **91**, 253–77.

Engel, E., Hayes, R. and Wang, X. (2007) The Sarbanes-Oxley Act and Firms Going Private Decisions, *Journal of Accounting and Economics*, **44**, 116–45.

Gao, F., Wu, J. and Zimmerman, J. (2009) Unintended Consequences of Granting Small Firms Exemptions from Securities Regulation: Evidence from the Sarbanes-Oxley Act, *Journal of Accounting Research*, **47**, 459–506.

Hall, B. (2002) The financing of research and development, *Oxford Review of Economic Policy*, **18**, 35–51.

Hamilton, J. (1994) *Time Series Analysis*, Princeton University Press: Princeton.

Hartman, T. (2007) The cost of being public in the era of Sarbanes-Oxley, Tech. rep., Foley & Lardner LLP.

Jain, P., Kim, J.-C. and Rezaee, Z. (2008) The Sarbanes-Oxley Act of 2002 and Market Liquidity, *The Financial Review*, **43**, 361–82.

Kalelkar, R. and Nwaeze, E. (2011) Sarbanes-Oxley Act and the quality of earnings and accruals: Market-based evidence, *Journal of Accounting and Public Policy*, **30**, 275–94.

Kamar, E., Karaca-Mandic, P. and Talley, E. (2007) Sarbanes-Oxleys Effects on Small Firms: What is the Evidence?, USC CLEO Research Paper C07-9, University of Southern California Law School.

Kang, Q., Liu, Q. and Qi, R. (2010) The Sarbanes-Oxley Act and Corporate Investment: A Structural Assessment, *Journal of Financial Economics*, **96**, 291–305.

Leuz, C., Triantis, A. and Wang, T. (2008) Why Do Firms Go Dark? Causes and Economic Consequences of Voluntary SEC Deregistrations, *Journal of Accounting and Economics*, **45**, 181–208.

Li, H., Pincus, M. and Rego, S. (2008) Market Reaction to Events Surrounding the Sarbanes-Oxley Act of 2002 and Earnings Management, *Journal of Law and Economics*, **51**, 111–34.

Linck, J., Netter, J. and Yang, T. (2009) The Effects and Unintended Consequences of the Sarbanes-Oxley Act on the Supply and Demand for Directors, *Review of Financial Studies*, **22**, 3287–328.

Litvak, K. (2007) Sarbanes-Oxley and the Cross-Listing Premium, *Michigan Law Review*, **105**, 1857–98.

Lysaght, M., Jaklenec, A. and Deweerd, E. (2008) Great Expectations: Private Sector Activity in Tissue Engineering, Regenerative Medicine, and Stem Cell Therapeutics, *Tissue Engineering: Part A*, **14**, 305–15.

Millar, J. and Bowen, B. (2011) Small and large firm regulatory costs: the case of the Sarbanes-Oxley Act, *Corporate Governance*, **11**, 161–70.

Paunov, C. (2012) The global crisis and firms' investments in innovation, *Research Policy*, **41**, 24–35.

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org.

Salter, B. (2005) The global politics of human embryonic stem cell science, Global Biopolitics Research Group Working Paper 2, University of East Anglia.

Sarbanes-Oxley (2002) Sarbanes-Oxley Act of 2002, Public Law 107-204, 107th Cong., 30 July 2002.

Singer, Z. and You, H. (2011) The Effect of Section 404 of the Sarbanes-Oxley Act on Earnings Quality, *Journal of Accounting, Auditing & Finance*, **26**, 556–89.

United States White House (2001) Embryonic Stem Cell Research. White House Fact Sheet., Downloaded from http://georgewbush-whitehouse.archives.gov/news/releases/2001/08/20010809-1.html.

USPTO (2013) USPTO Patent Full-Text and Image Database, downloaded from http://patft.uspto.gov/.

Wintoki, M. (2007) Corporate Boards and Regulation: The Effect of the Sarbanes-Oxley Act and the Exchange Listing Requirements on Firm Value, *Journal of Corporate Finance*, **13**, 229–50.

Zhang, I. (2007) Economic consequences of the Sarbanes-Oxley Act of 2002, *Journal of Accounting and Economics*, **44**, 74–115.

Zucchini, W. and MacDonald, I. (2009) *Hidden Markov Models for Time Series: An Introduction Using R*, CRC Press: Florida.