# Idle Oil Wells:

# Half Empty or Half Full?

Lucija Muehlenbachs *

March 25, 2009

**Abstract**

There are hundreds of thousands of oil and gas wells across North America that are not currently producing oil or gas. Many of these wells have not been permanently decommissioned to meet environmental standards for permanent closure, but left in a inactive state that enables the well to be more easily reactivated. Some of the wells have been temporarily inactive for more than sixty years which begs the question of whether these wells are inactive because of the option value of reactivation or because of the hefty cost of complying with environmental regulations for permanent closure. The goal of this paper is to determine what changes in prices, technology or policy would be needed to see an increase in reactivated or decommissioned wells. Historical data of production decisions from 84 thousand wells and expected recoverable reserves from 47 thousand pools are used to estimate a dynamic discrete choice model of operating state. The estimated structural model then allows us to examine what conditions might push any of the inactive wells out of the hysteresis in which they reside.

*Jel–Classification:* C61, Q32, Q41

*Keywords:*dynamic structural estimation; environmental remediation; oil and gas

*I thank my advisors John Rust and Marc Nerlove. I also thank the Alberta Energy and Resources Conservation Board for access to the data, and the Chicago-Argonne Institute on Computational Economics for the optimization software. Department of Agricultural and Resource Economics, University of Maryland, College Park, USA; email *lmuehlenbachs@arec.umd.edu*

# 1 Introduction

Hundreds of thousands of oil and gas wells scattered across North America are currently not producing oil or gas. Once a well has reached the end of its productive life it must be decommissioned to prevent contamination of groundwater and to reclaim the surrounding land. On the other hand, as prices rise and technology emerges, reserves that were once unprofitable to extract become profitable and inactive wells are reactivated. To maintain the potential for future production, regulators of the industry allow non-producing wells to remain inactive, and do not force producers to decommission them permanently. Many wells, however, are dry holes or are sufficiently damaged that drilling a new well would be less expensive than putting it back into production. The trade-off of not mandating abandonment is that the inactive wells carry the risk of explosion, lost productivity and contamination of freshwater, vegetation and soil [Kubichek et al., 1997, Williams et al., 2000]. There is also the risk that the owners of the inactive wells will declare bankruptcy before undertaking the expenses of abandonment and environmental remediation. Texas for example, has roughly 10,000 orphaned wells for this reason[RRC, 2006].

Regulators face a fine balancing act of not impeding extraction while ensuring environmental obligations are met. Any regulation to combat the number of inactive wells is made under incomplete information (on the condition of the well and oil operator's true intentions) and under uncertainty (in future technology and prices). In Alberta, Canada (the source of the data used in this paper) there are wells that have been inactive for over sixty years, some of which may have a viable future, while some of which may be inactive as a means of evading the environmental clean up costs. Before implementing policies to deal with the accumulation of inactive wells (either policies to induce production or decommissioning), it is important to know how the number of inactive wells would change exogenously with prices or technology. The goal of this paper to examine which changes (in prices, recovery technology, and finally, policy) would be necessary to promote a significant decrease in the percentage of inactive wells. To this end, a dynamic discrete choice model of the optimal operating state (active, inactive, or abandoned) is developed. The parameters of the structural model are estimated using data from 84 thousand Albertan oil and gas wells from 2000 to 2007. This decision depends on many observable and unobservable states of nature, as well as a producer's perception of what the future state of nature will be. To account for the effect of changes in expected recovery because of technological improvements, discoveries or the reevaluation of reserves in place, this paper makes use of a panel of official reserve estimates for all non-confidential pools in Alberta (47 thousand pools).

## 1.1 Literature Review

This paper uses dynamic programming and data on decisions under uncertainty to explore the ideas behind the literature on real options. Much of the literature on real options relies on examples from the natural resource industry, so much so that models for most of the discrete decisions facing firms in the industry have been developed. The same three choices presented in this paper (to activate, inactivate or decommission a project) have already been modeled by Brennan and Schwartz [1985], Dixit and Pindyck [1994], Gamba and Tesser [2007]. Notably, these authors did not apply their models to real data. Throughout this literature however the permanent closure option is often downplayed. Abandonment costs are treated as negligible or null[Brennan and Schwartz, 1985, Dixit and Pindyck, 1994], or the option of final abandonment is completely left out of the choice set[Moel and Tufano, 2002, Slade, 2001, Mason, 2001, Paddock et al., 1988]. By assuming away abandonment costs, the previous literature has overlooked the case of firms that continue to pay for the option to reactivate a project, even when they have no intention to reactivate. This is a concern for any industry where abandonment costs are high, which is especially true in the natural resources industry due to the large imprint that these projects leave on the environment. If a firm "temporarily" closes a site, when in fact there is no potential for reactivation, regulators have reason to implement policies to ensure that environmental obligations are met.

To implement a real options model requires either cost data or estimates of cost. Cost data is difficult to obtain and investigators have dealt with this requirement in different ways. The first has been to forgo the need for cost data by illustrating real options by way of a numerical example [Mason, 2001, Brennan and Schwartz, 1985, Dixit and Pindyck, 1994]. Those that have used cost estimates have been restricted to small sample sizes. Using the U.S. Geological Survey's cost estimates of development and exploration for offshore petroleum leases, Paddock et al. [1988] evaluate a real options model with data on 21 offshore petroleum tracts. Moel and Tufano [2002] examines mine opening and closings using average annual cash costs (including fixed and marginal costs) of 285 mines, with no data on closing and opening costs. Slade [2001] is able to evaluate mine opening and closings using panel data on average production costs for 175 mines (but again without data on closing and opening costs). The need for cost estimates before examining the real options models has confined these authors to small datasets. Because data on production is easier to obtain than data on costs (production reporting is usually mandatory, whereas cost data is proprietary) the gains of not being reliant on cost data is therefore massive. In this paper, costs are estimated using data on production decisions via Rust [1987]'s Nested Fixed Point Algorithm. To estimate the costs using the operating decisions is a powerful way to take advantage of the extensive production dataset available.

The model developed here can also contribute to the literature on financial bonding to ensure environmental remediation. One of the main reasons for a policy to induce prompt environmental clean up is the risk that the firm will declare bankruptcy. The concern that oil and gas companies may walk away from their environmental obligations has been brought up by Boyd [2001], Parente et al. [2006], Ferreira et al. [2003]. While these papers discuss bonding mechanisms, the model here can be used to quantify the effect of a bond on production as well as the choice to undertake cleanup.

# 2 Background

The vast number of oil and gas wells in Alberta is no indicator of the quantity of remaining reserves under those wells. The remaining reserves of conventional oil (that which is pumped out of the ground from a well) is eclipsed by the potential contribution from non-conventional oil (also known as bitumen) that is produced from the oil sands. The remaining established reserves of conventional oil in Alberta were estimated in 2007 at 1.5 billion barrels, compared to 173 billion barrels of unconventional oil[ERCB, 2008]. A further illustration of the size of Alberta's conventional reserves is by comparing production from Alberta's wells to global oil and gas production. In 2007, Alberta's conventional oil production contributed 0.52 million barrels a day, or .6% of the world's average oil supply of 84 million barrels a day[EIA, 2009]. Alberta's natural gas production contributed .135 trillion $m^3$, or 4.6% of the world's total 2.942 trillion $m^3$ [1] in 2006[EIA, 2008].

Producers leaving a well inactive commonly claim that they are waiting for the price to increase to a point at which the well will become profitable. While a well might be drilled in a pool with a high estimated amount of oil in place, the current recovery factor, or fraction of the oil in place that can be recovered is likely to be low (in Alberta the average percent recoverable is only 26%). Therefore, an operator might want to hold on to a well that taps into a massive reserve, even if it is not currently economical to produce oil or gas, in the hopes that prices or technology improve. Another reason there are so many inactive wells caught in a state of hysteresis because of the high costs of either retrofitting the well for production or plugging and abandoning the well. The preceding reasons are directly modeled in this paper. The following are some reasons for inactivity that are not explicitly entered into the model (but only through an error term that compensates for unobservable states): (1) technical difficulties (for example, blockage in the wellbore, a leak caused by corrosion or erosion, an external fire, or a temperature change causing mechanical failure), (2) pipeline failure or pipeline capacity reached, (3) a mandated suspension for exceeding the maximum rate limit[2] assigned to the well by the regulator, or (4) a change in

[1] equivalent to 103.9 trillion cubic feet

[2] Various wells must conform to maximum rate limitations set by the industry regulator. These limits are to

ownership of the well.

The development of enhanced recovery methods has brought wells back into production after many years of inactivity. These extraction methods have improved over the last century, with the introduction of horizontal wells, progressive cavity pumps, cyclic steam stimulation, and coiled tubing[Beliveau and Baker, 2003]. Today most operators use enhanced recovery methods, most commonly by injecting water or gas into the well or a nearby well. There are many other techniques to increase the production rate, such as generating carbon dioxide in-situ, dissolving minerals with acid, injecting hot fluid or steam, or creating in-situ combustion. There are also various pumps that provide mechanical energy such as sucker rod pumping, electrical submersible pumping, hydraulic piston pumping, hydraulic jet pumping, progressive cavity pumping, and plunger lift systems.

With the introduction of these methods, recoverable reserves have been seen to increase, rather than decrease with time. For example from 1978 to 1993 the decrease in U.S. natural gas reserves by production was approximately the same as the increase in reserves, where 87% of the increase came from existing fields [Beliveau and Baker, 2003]. Reserve growth was first examined by Arrington [1960] using his own company's reservoir data. And since then reserve growth has been studied using state or state subdivision estimates of initial established reserves from the American Petroleum Institute[Morehouse, 1997], or a small number of pools[Verma and Henry, 2004].

But on the other hand without proper plugging and abandonment (and in some cases, even after proper plugging and abandonment) a well poses a risk to vegetation, soil, surface water, and underground aquifers. The wellbore can extend thousands of meters underground, and once drilled, it is only the steel casing, and often cement that isolates the different formations. The casing might rust out or crack (especially when lifting sand or saltwater along with the hydrocarbons) and contaminants such as uranium, lead, salt, iron, selenium, sulfates, and radon [Kubichek et al., 1997] may enter into freshwater.

The most prevalent contaminant from oil and gas wells into fresh water aquifers has been

ensure that the cumulative amount of oil or gas extracted is maximized. In our model we will not truncate by maximum rate limit because only about 10 percent of the wells have limits placed on them, and for only a portion of those wells is the rate limit binding.

methane gas[3]. The concern surrounding methane migration is that high concentrations can lead to explosions (there are instances of homes and windmills exploding). The presence of methane also foreshadows worse events to come. Methane indicates that a leak exists, meaning that other contaminants that are in lower concentrations, and slower moving than methane, follow shortly behind. Those more hazardous contaminants are the carcinogens uranium and radon, the probable carcinogen selenium, and the neurotoxin lead.

In Alberta, there is no time limit for leaving a well inactive; it is up to the firm to decide when to permanently abandon their well. In 2004 the ERCB issued a directive [ERCB013, 2004] to increase requirements for inactive wells, whereby low risk wells (or those wells that have noncritical levels of sour gas $H_2S > 5\%$, or gas wells with an open flow potential of less than 28 thousand $m^3$/day) were to be inspected every 1-5 years, and medium and high risk wells (non-flowing oil wells with sour gas) must either place a packer and a tubing plug, or a bridge plug in the wellbore where high risk (critical level of sour gas) the bridge plug is capped with 8 meters of cement. Low risk wells . Beginning in December 2007, low risk wells that were inactive for over 10 years must be suspended according to the requirements for medium risk wells. In 2002 Alberta implemented a bonding scheme where companies whose liabilities exceeded their assets have to post a bond for the difference. The abandonment costs used in the calculation range from $7,200 to $90,000. Land reclamation costs are also included which vary by location from $13,200 (grasslands area east) to $33,700 (alpine area). However, abandonment and reclamation costs may drastically surpass these figures. For example, the Orphan Well Association has previously spent $800,000 to abandon a single well, and over $1,000,000 to remediate the land around another[Orphan, 2006]. If the bond is less than the abandonment cost, there is no incentive for the firm to abandon the well. Furthermore, it is likely that the liability figures used are underestimates because they are calculated from figures submitted in a voluntary survey of costs. If a company becomes defunct, the Orphan Well Association will collect the required abandonment and reclamation costs from the remaining firms in the industry based on each firms share of industry liability.

[3] It is possible that methane may enter groundwater through shallow reservoirs that the water well penetrates, swamps or landfills, thereby relinquishing obligation by the oil and gas producer. To determine the source of the methane gas isotopes are used.

A royalty on production must be paid to the owner of the mineral rights. The royalty structure in Alberta adjusts according to price, when the pool was discovered, and productivity. When the mineral rights are not owned by the government, a royalty is paid to the owners of the mineral rights, and a freehold mineral tax is paid to the government. In Canada natural resources are owned by the province. Approximately 81 percent of oil and gas is owned by the Crown, and the remaining is owned by the Federal government (in National Parks), by First Nations, and by Freehold land owners where the title had been given to the railway companies, the Hudson's Bay Company, and the early settlers Energy [2007]. Up until January 2009, the royalty formulas remained the same since 1993[MineralAct, 2008], coinciding with the study period and strengthening the claim of a consistent regulatory regime.

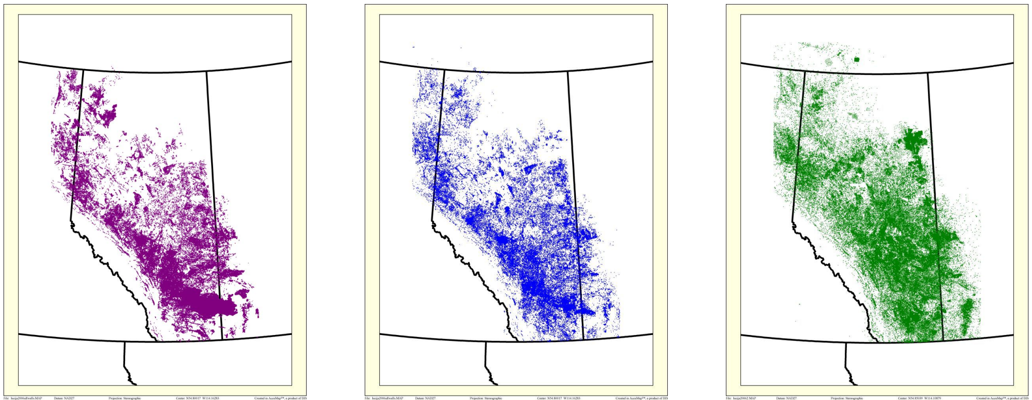
## 3 Data

The following five datasets of the Albertan oil and gas industry are used. The first is a panel dataset of production from all oil and gas wells in Alberta. Obtained through IHS Incorporated (who distribute the records collected by the industry regulator, the Alberta Energy Resources Conservation Board (ERCB)), this dataset contains monthly oil and gas production dating back to 1924. For over 340 thousand wells there is information on the location (latitude and longitude as well as the name of the field and pool that the well is on), depth, spud date (when the drill hit the ground), licence date, on production date, name of the current operator and name of the original operator (unfortunately there is no information on whether the well switched hands other than between the original and current operators).

The second dataset is a panel of official reserve estimates of all nonconfidential pools in Alberta from both the ERCB and the National Energy Board of Canada. The dataset was kindly received from the ERCB and contains roughly 33 thousand to 56 thousand oil and gas pools per year from 2000 to 2007. The date that the estimate was last reviewed is listed and therefore the data can be extended to include estimates starting from 1989. Information on reserves are broken down into: 1)initial oil or gas in place 2) recovery factor 3) initial established reserves (equal to oil or gas in place multiplied by the recovery factor) 4) remaining established reserves (initial established reserves less cumulative production and surface loss). The reserve dataset also contains information on characteristics of the pools such as porosity, initial pressure, area, density, thickness, gas saturation (fraction of pore space in the reservoir occupied by gas) etc.

The third dataset is of wells that were permanently abandoned. For all wells that were surface abandoned, there is data on the month and year of abandonment. There is also information on all wells that either received a reclamation certificate or were exempt from undergoing reclamation. Certification of the reclamation of surrounding land is in the jurisdiction of Alberta Environment, whereas abandonment (downhole sealing of the wellbore) is in the jurisdiction of the ERCB.

The Petroleum Services Association of Canada (PSAC) has divided Alberta into 7 areas



(a) Active Wells (b) Inactive Wells (c) Abandoned Wells

Figure 1: Oil and Gas wells in Alberta in 2007

based on similar costs in production and drilling. The shape files that delineate the PSAC cost areas by latitude and longitude were entered into ArcView GIS to allocate each well in the production dataset into its given PSAC area. And the final dataset is the average wellhead price of crude oil and natural gas in Alberta, obtained from the Canadian Association of Petroleum Producers' Statistical Handbook [4]. The wellhead price is inflated to 2007 dollars using Statistics Canada's quarterly machinery and equipment price index for mining, quarries and oil wells.

Excluding coalbed methane, heavy oil, injection and water wells there are 341,136 wells in the production dataset. Of these wells, 94,365 are found in a pool that is listed in the reserves dataset. The data is then divided into oil wells and gas wells so that separate dynamic programming models can be made. Dropping observations of wells that traverse both oil and gas pools does not further reduce the size of the dataset significantly. After deleting all wells that produce oil and gas from different pools, the final sample contains 84,155 wells distinguished as either oil wells or gas wells.

A panel is created where each well is classified as either active, inactive, or surface abandoned for each year starting with the year the well was drilled until 2007. A well is classified as active if it produced any amount of gas or oil within that year, classified as inactive if it has not produced oil or gas in twelve months or more, and classified as abandoned according to the dataset of

[4] http://www.capp.ca/default.asp?V_DOC_ID=1072&SectionID=5&SortString=TableNo

Table 1: Distribution of wells according to status within age groups

| Age (in years) | Number of Observations | | Proportion Active | | Proportion Inactive | | Proportion Abandoned | |
|---|---|---|---|---|---|---|---|---|
| | (Subsample) | (Full) | (Subsample) | (Full) | (Subsample) | (Full) | (Subsample) | (Full) |
| $0 \leq age < 10$ | 240314 | 1661127 | 0.72 | 0.50 | 0.25 | 0.29 | 0.04 | 0.20 |
| $10 \leq age < 20$ | 152717 | 1046268 | 0.50 | 0.37 | 0.33 | 0.26 | 0.17 | 0.37 |
| $20 \leq age < 30$ | 106496 | 727899 | 0.50 | 0.34 | 0.26 | 0.22 | 0.25 | 0.44 |
| $30 \leq age < 40$ | 44018 | 404664 | 0.49 | 0.25 | 0.26 | 0.29 | 0.26 | 0.46 |
| $40 \leq age < 50$ | 41773 | 327113 | 0.50 | 0.17 | 0.23 | 0.45 | 0.27 | 0.38 |
| $50 \leq age < 60$ | 14487 | 138503 | 0.32 | 0.10 | 0.21 | 0.53 | 0.47 | 0.37 |
| $60 \leq age < 70$ | 2090 | 11655 | 0.37 | 0.15 | 0.36 | 0.46 | 0.27 | 0.39 |
| $70 \leq age < 80$ | 984 | 5201 | 0.35 | 0.15 | 0.24 | 0.48 | 0.41 | 0.37 |

Notes: Data from 1989-2007. The "full" dataset comprises all wells in Alberta. The "subsample" dataset contains only the wells that produce from a single reservoir, and that have pool information.

abandonments obtained from the ERCB. A snapshot of wells in different operating states in 2007 is found in Figure 1.

Table 1 shows the distribution of different operating states according to age groups for the full dataset of all wells in Alberta, as well as the subsample used in this paper. The largest divergence between the full dataset and the subsample is the abandonment rate of wells less than 10 years old. According to the subsample (that is wells that have pool information), new wells are not being abandoned, whereas in the full dataset 20% of new wells are abandoned. This is expected in that if there is any wildcat well drilled but found not to be associated with any pool, it is more likely to be abandoned, but it will not appear in the subsample (as the subsample has only wells with pool information).

Summary statistics for the data is found in Table 2. There is a very long tail in the distribution of reserve size; the majority of the observations are wells with small reserves while a small minority have enormous reserves.

Table 3 shows a multinomial logit regression for the discrete choice of active, inactive or abandon. As expected, the older the well the more likely it is to be abandoned, and the more likely it is to be inactive than active. Also as expected, when there are large remaining reserves the wells are less likely to be abandoned, and more likely to be active than inactive. For oil and gas wells, it is as expected that with higher prices wells are less likely to be abandoned, however not as expected, under high gas prices, gas wells are more likely to be inactive than active. The rest of the regressors help guide what variables the data should be clustered on.

Table 2: Summary Statistics

| Variable | No. of Obs. | Mean | Std. | Min | Max | Unit |
|---|---|---|---|---|---|---|
| $\overline{Q}_g$ | 138229 | 15.88 | 62.87 | 0 | 8800 | $E^6m^3$ |
| $\overline{Q}_o$ | 51490 | 29.09 | 235.39 | 0 | 43871.43 | $E^3Barrels$ |
| Price$_{\text{g}}$ | 19 | 126.12 | 77.32 | 53.94 | 293.90 | $Dollars/E^3m^3$ |
| Price$_{\text{o}}$ | 19 | 32.61 | 13.50 | 16.10 | 64.45 | $Dollars/Barrel$ |
| $Q_g$ | 138229 | 37.24 | 362.05 | 0 | 51271.00 | $E^6m^3$ |
| $Q_o$ | 51490 | 164.18 | 2302.00 | 0 | 367199.14 | $E^3$ Barrels |
| No. of wells in Pool$_{\text{g}}$ | 164457 | 3.66 | 34.31 | 1 | 4153 | Wells |
| No. of Firms in Pool$_{\text{o}}$ | 164457 | 1.26 | 1.14 | 1 | 87 | Firms |
| $q_g$ (Subsample) | 285019 | 1.89 | 8.75 | 0.0001 | 568.39 | $E^6m^3$ |
| $q_o$ (Subsample) | 153182 | 7.44 | 18.08 | 0.000629 | 3428.87 | $E^3Barrels$ |
| $q_g$ (Full) | 1470435 | 1.88 | 8.37 | 0.0001 | 930.39 | $E^6m^3$ |
| $q_o$ (Full) | 618297 | 10.68 | 25.33 | 0.000629 | 3428.87 | $E^3Barrels$ |
| Depth | 83668 | 1267.83 | 699.33 | 90.9 | 6552 | $m$ |
| Porosity$_{\text{g}}$ | 21767 | 0.20 | 0.07 | 0.01 | 0.4 | Fraction |
| Porosity$_{\text{o}}$ | 24967 | 0.16 | 0.07 | 0.01 | 0.36 | Fraction |
| Density$_{\text{g}}$ | 21767 | 0.64 | 0.08 | 0.54 | 2.03 | Fraction |
| Density$_{\text{o}}$ | 24967 | 866.79 | 45.50 | 708 | 999 | $kg/m^3$ |
| Initial Pressure$_{\text{g}}$ | 21767 | 9130.55 | 7666.57 | 130.00 | 99625 | kPa |
| Initial Pressure$_{\text{o}}$ | 24967 | 12597.58 | 5707.25 | 1442 | 61097 | kPa |
| Compressibility$_{\text{g}}$ | 21767 | 0.89 | 0.08 | 0.43 | 8.88 | Fraction |
| Temperature$_{\text{o}}$ | 24967 | 50.56 | 19.95 | 9 | 350 | $C^\circ$ |
| Water Saturation$_{\text{o}}$ | 24967 | 0.32 | 0.12 | 0.06 | 0.82 | Fraction |
| Area of Pool$_{\text{o}}$ | 24967 | 259.14 | 956.31 | 2.47 | 44207.15 | Acres |
| Area of Pool$_{\text{g}}$ | 21767 | 1482.74 | 12358.72 | 2.47 | 546379.65 | Acres |
| Well Spacing$_{\text{g}}$ | 138229 | 589.84 | 3831.68 | 0.02 | 392564 | Acres/Well |
| Well Spacing$_{\text{o}}$ | 51490 | 84.43 | 140.88 | 0.10 | 19521.33 | Acres/Well |
| Wells per Firm (Subsample) | 1103 | 316.38 | 2195.36 | 1 | 47720 | Wells |
| Wells per Firm (Full) | 2823 | 128.96 | 1379.09 | 1 | 47720 | Wells |
| Pool Discovery Year$_{\text{g}}$ | 21767 | 1989 | 13 | 1904 | 2007 | Year |
| Pool Discovery Year$_{\text{o}}$ | 24967 | 1988 | 12 | 1910 | 2006 | Year |

Notes: "Full" encompasses all oil (o) and gas (g) wells in Alberta and "subsample" encompasses only wells that have pool information. Data on remaining reserves ($Q$) is listed for pools 1989-2007. Extraction ($q$) is listed for wells, 1989-2007. The pool specific variables, depth, porosity, density, initial pressure, compressibility, temperature, water saturation, area of pool and discovery year are time invariant in the data. Data on the number of wells held by a firm is a snapshot of 2007.

Table 3: Mulitnomial Logit Estimates

| Oil Wells | | | Gas Wells | | |
|---|---|---|---|---|---|
| | Activate vs. Abandon | Inactivate vs. Abandon | | Activate vs. Abandon | Inactivate vs. Abandon |
| Intercept | 1.13** | -0.76** | Intercept | 1.18** | -0.23 |
| | -0.21 | (-0.22) | | (0.20) | (0.22) |
| $\overline{Q}_o$ | 0.18** | 0.17** | $\overline{Q}_g$ | 0.0078** | 0.0075** |
| | (3.07E-03) | (3.08E-03) | | (4.23E-04) | (4.25E-04) |
| $\text{Price}_o$ | 0.0043** | 0.0083** | $\text{Price}_g$ | 3.04E-06** | 1.13E-06** |
| | (5.38E-04) | (5.52E-04) | | (1.27E-07) | (1.32E-07) |
| $\text{Porosity}_o$ | 1.87** | 1.47** | $\text{Porosity}_g$ | -1.04** | 1.69** |
| | (1.67E-01) | (1.70E-01) | | (0.15) | (0.16) |
| $\text{Initial Pressure}_o$ | 3.44E-05** | 2.76E-06 | $\text{Initial Pressure}_g$ | 1.98E-05** | 2.06E-05** |
| | (3.18E-06) | (3.26E-06) | | (2.84E-06) | (2.96E-06) |
| $\text{Depth}_o$ | -0.00025** | 1.5E-05 | $\text{Depth}_g$ | 3.45E-04** | 2.88E-04** |
| | (4.68E-05) | (4.72E-05) | | (2.46E-05) | (2.57E-05) |
| $\text{Density}_o$ | -0.00025 | 0.0012** | $\text{Density}_g$ | 0.89** | 1.61** |
| | (2.35E-04) | (2.41E-04) | | (0.18) | (0.18) |
| $\text{Age}_o$ | -0.069** | -0.043** | $\text{Age}_g$ | -0.083** | -0.051** |
| | (7.21E-04) | (6.94E-04) | | (6.31E-04) | (6.37E-04) |
| No. of Firms in $\text{Pool}_o$ | 0.00027 | -0.019** | No. of Firms in $\text{Pool}_g$ | 0.037** | 0.0075** |
| | (-0.0029) | (3.05E-03) | | (0.0026) | (2.85E-03) |
| $\text{Over}_o$ | -0.23** | -0.071** | $\text{Over}_g$ | -0.074** | -0.012 |
| | (2.81E-02) | (2.90E-02) | | (0.016) | (0.017) |
| Wells per Firm | -8.64E-06** | -6.00E-06** | Wells per Firm | 4.06E-06** | 8.50E-08 |
| | (6.88E-07) | (7.05E-07) | | (6.04E-07) | (6.41E-07) |
| No. Wells in $\text{Pool}_o$ | 0.0023** | 0.00089** | No. Wells in $\text{Pool}_g$ | 0.0018** | 6.47E-04** |
| | (6.94E-05) | (7.47E-05) | | (1.37E-04) | (1.52E-04) |
| Water $\text{Saturation}_o$ | -0.85** | -1.78** | | | |
| | (7.45E-02) | (7.71E-02) | | | |
| Gas to Oil Ratio | 0.0062** | 0.0034** | | | |
| | (2.97E-04) | (3.06E-04) | | | |
| $\text{Temperature}_o$ | 0.0042** | 0.010** | | | |
| | (9.77E-04) | (9.86E-04) | | | |
| | | | $\text{Compressibility}_g$ | -0.1 | -0.32** |
| | | | | (0.15) | (0.16) |
| Log Likelihood | -164277.07 | | | -173110.89 | |
| No. of Observations | 171923 | | | 196242 | |

Abandonment is the reference choice in the logit. Over is a dummy variable for whether the well spacing limit has been reached. ** parameter estimates are significantly different from zero at the 2.5% level. Standard errors are in parentheses.

# 4 Model

## 4.1 Deterministic Model

A dynamic programming framework lends itself to the modeling of the optimal operating state of an oil well–a dynamic decision made under much uncertainty. The current choice of operating state depends on the future stream of profits from that choice which will be a result of uncertain prices, technology and reserves. The operator's problem can be modeled as an infinite time Markov Decision Process, which as explained in [Rust, 1994], includes a decision variable $d$ (that is here to extract, 1, temporarily stop extraction, 2, or to permanently abandon and remediate environmental damages, 3). It is assumed that the operator is rational, and follows a decision rule $d_t = \delta_t(s_t)_{t=0}^{\infty}$ for the optimal choice under all possible states of nature, $s_t$. The decision rule maximizes the expected discounted sum of profits, $V_t(s_t) = \max_{\delta} \mathrm{E}_{\delta}\left[\sum_{t=0}^{\infty} \beta^t \pi_t | s_t\right]$, where $V_t(s_t)$ is the value function for the well when following the optimal decision, $\beta$ is the discount factor, $0 \leq \beta \leq 1$, and $\pi(\cdot)$ is the instantaneous profit from the well. The vector of state variables, $s$, include the age of the well, $A$, the wellhead price of the hydrocarbon, $P$, the remaining reserves per well[5], $\overline{Q} = Q/n$, and the current operating state $o$. The current operating state ($o$= active, 1, inactive, 2, or abandoned, 3) is endogenous to the decision, and the remaining reserves per well changes both endogenously (when extracting) and exogenously (upon revisions of reserve estimates or if another well in the same pool extracts).

If the operator decides to extract, it will receive the current wellhead price, $P$, per quantity, $q$, extracted, less corporate income tax $\tau$, royalties $R$, lifting costs $C$, and a fixed cost, $M_1$. Leaving the well inactive also entails an annual fixed cost, $M_2$. If the operator decides to permanently abandon the well, they do not have to pay the annual fixed costs, although they also forgo any future extraction. The profit from a single period is:

$$\pi(s,d) = \begin{cases} \left((1-R)P - C\right)q - M_1 - \tau \max\{\left((1-R)P - C\right)q - M_1, 0\} - SC_{(o\to 1)} & \text{if d=1} \\ -M_2 - SC_{(o\to 2)} & \text{if d=2} \\ -SC_{(o\to 3)} & \text{if d=3} \end{cases} \tag{1}$$

The lifting cost $C(P, \overline{Q}, A)$ is increasing in price and age, and decreasing in per well reserves[6].

[5]Livernois [1987] also incorporates the concept of reserves per well. In the case of Alberta, as can be seen in table 2 of summary statistics, there are as many as 4153 wells in a pool, hence the division of reserves by the number of wells.

[6]Chermak and Patrick [1995] and Foss and Gordon [2007] show how the lifting cost of natural gas depends on quantity extracted, and remaining reserves. Chermak and Patrick [1995] use data from 29 gas wells in Wyoming and Texas from 1988 to 1990, and Foss and Gordon [2007] uses data from 22 gas wells in Alberta for roughly 3

As costs will also depend on well depth, porosity, location etc., the dynamic program will be solved for data that is clustered on these time invariant states. The royalty rate $R(P, q, A)$ adjusts according to price, when the pool was discovered, and quantity produced. $\tau$ is the corporate income tax and is assumed flat for all wells.

Switching costs, $SC$, are incurred when the operator moves between operating states. The operator pays $SC_{(1\to2)}$ to temporarily inactivate an active well, $SC_{(2\to1)}$ to reactivate an inactive well and $SC_{(1,2\to3)}$, to decommission a well (assumed to be the same for an active or inactive well). Leaving the well in its current state entails no switching costs, $SC_{(1\to1)} = 0$, $SC_{(2\to2)} = 0$, and $SC_{(3\to3)} = 0$. Abandonment is an absorbing state (in the sample there are only 261 observations of a switch from abandoned to active, where as there were 22308 observed inactivations, 15369 reactivations of inactive wells, 1917 active wells abandoned, and 3664 inactive wells abandoned).

The expected present discounted value of the well can be expressed as the unique solution to the Bellman equation:

$$V(s) = \max_d [\mathrm{E}\pi + \beta \mathrm{E}V(s)]$$

Where $\mathrm{E}V(s)$ is the conditional expectation of the value function in future state $s'$:

$$\mathrm{E}(V(s)) = \int_{s'} V(s')g(s'|s)\mathrm{d}s'$$

The transition of $s$ to $s'$ is assumed to be a Markov process with a transition probability density function $g(s'|s)$. This density contains that of price and recoverable reserves as follows. Price is assumed to follow the exogenous process $f_P(P'|P, \theta_P)$, characterized by parameters $\theta_P$. The transition of recoverable reserves depends on whether the well is the only well in the pool or not. Quantity extracted, $q$, is modeled as a random draw from a density that depends on the reservoir size as well as age of the well, $f_q(q|\overline{Q}, A, \theta_q)$, so therefore recoverable reserves decrease with extraction. For single-well pools, when the decision is to extract, next period reserves are $\overline{Q}' = \overline{Q} - \int_0^Q q f_q(y|\overline{Q}, A, \theta_q)dy$, while reserves in multi-well pools decrease by $\int_0^Q q f_q(y|\overline{Q}, A, \theta_q)dy$ in every period. $\overline{Q}$ can also increase or decrease because of reassessment, improved technology, discoveries or additions. Therefore exogenous to the decision, reserve size also follows another process $f_{\overline{Q}}(\overline{Q}'|\overline{Q}, \theta_{\overline{Q}})$. $\overline{Q}$ also decreases whenever another well is drilled in the pool. A simplification

years. They both find that operating costs increase with quantity extracted, and decrease with remaining reserves. It is expected that extraction costs rise as reserves are depleted, however, Livernois and Uhler [1987] explain that the discovery of new reserves can increase the reserves by more than what is extracted, but these new reserves are more costly to extract. This is how Livernois and Uhler [1987] explain a positive relationship between extraction costs and reserves using aggregate data from the Albertan oil industry. However, upon disaggregation, they find the typical results of extraction costs increasing with reserve depletion, and quantity extracted.

for the time being includes this probability of decrease in the probability density $f_{\overline{Q}}$.

### 4.2 Estimation Strategy: Introduction of Unobservables

The decision rule that maximizes the Bellman equation as specified above dictates exactly what the operator should do under any given age, price and recoverable reserves. But for any decision rule, the data will not perfectly coincide because, in reality, there are many other factors that will determine the operator's decision. Therefore, following Rust [1987] we add an element, $\epsilon_d$, to the profit under each alternative that is observed by the operator but not by the econometrician. This unobserved component accounts for unobserved heterogeneity in the firms or wells, or more specific characteristics that influence costs such as pressure, condition of the casing etc. We use Rust's assumptions on $\epsilon_d$ in order to facilitate estimation. First, $\epsilon_d$ enters the profit in an additively separable (AS) way:

$$\pi(s,d) = \begin{cases} ((1-R)P - C)\,\mathrm{E}q - M_1 & \text{if d=1} \\ \qquad -\tau \max\{((1-R)P - C)\,\mathrm{E}q - M_1, 0\} - SC_{(o\to 1)} + \epsilon_1 & \\ -M_2 - SC_{(o\to 2)} + \epsilon_2 & \text{if d=2} \\ -SC_{(o\to 3)} + \epsilon_3 & \text{if d=3} \end{cases} \tag{2}$$

Where $\mathrm{E}q = \int_0^Q q f_q(y|\overline{Q}, A, \theta_q)dy$. Also following [Rust, 1987, 1988], the transition probability for next period state variables can be factored:

$$g(s', \epsilon'|s, \epsilon, d) = f(s'|s, d)\rho(\epsilon'|s')$$

. Or specifically in the case of the model:

$$g(\overline{Q}', P', \epsilon'|\overline{Q}, P, \epsilon, d) = f_q(\overline{Q} - \overline{Q}'|\overline{Q}, d) f_{\overline{Q}}(\overline{Q}'|\overline{Q}) f_P(P'|P)\rho(\epsilon'|\overline{Q}', P')$$

. $\rho(\cdot)$ is the transition probability density function of $\epsilon$ the unobserved state variable. This factorization is the Conditional Independence (CI) assumption that implies that $\epsilon'$ entirely depends on the observed variables, and not on $\epsilon$, while the transition of the observed state variables entirely depends on the current observed states and not on $\epsilon$. This assumption allows for $\epsilon$ to be integrated out from the Bellman equation, allowing for $EV$ to no longer be a function of $\epsilon$.

Under the additional assumption that $\epsilon$ is independent and identically Type I Extreme Value distributed the Bellman equation becomes:

$$V_\theta(s,\epsilon) = \max_d [v_\theta(s,d) + b\epsilon(d)]$$

where $v_\theta$ is the fixed point of $v_\theta = \Gamma(v_\theta)$.

And $\Gamma_\theta$ is a contraction mapping:

$$\Gamma_\theta(v)(s,d) = \mathrm{E}(\pi(s,d,\theta)) + \beta \int_{s'} b \log \sum_{d'=1}^{3} \left[ \exp\left\{ \frac{v_\theta(s',d')}{b} \right\} \right] f(s'|s,d)\mathrm{d}s' \tag{3}$$

$b$ is the standard deviation of $\epsilon$, or the scale parameter from the Extreme Value distribution. When $b$ approaches zero, $V_\theta(s,\epsilon)$ converges to the ordinary Bellman equation. Computing the fixed point of $\Gamma_\theta$ is far easier than had there not been the CI assumption, in which case we would have to solve for the fixed point $V_\theta$ to Bellman's equation:

$$V_\theta(P,Q,\epsilon) = \max_d [\mathrm{E}(\pi(s,d,\theta)) + b\epsilon(d) + \beta \mathrm{E} V_\theta(s,\epsilon)]$$

Because of the CI assumption, $\mathrm{E}V_\theta$ is not a function of $\epsilon$ and therefore we do not have to integrate out the $\epsilon$ distribution to obtain the choice probabilities for the likelihood function. Also, the assumption of the Extreme Value distribution allows for a closed form solution for the choice probabilities–that of the multinomial logit:

$$p(d|s,\theta) = \frac{\exp \frac{v_\theta(s,d)}{b}}{\sum_{d'} \exp \frac{v_\theta(s,d')}{b}} \tag{4}$$

The likelihood of all of the observed operating decisions and state variable transitions as a function of $\theta$ is:

$$L_f(\theta) = \prod_{i=1}^{N_i} \prod_{t=1}^{T_i} p(d_t^i|s_t^i\theta) f(s_{t+1}^i|s_t^i,\theta)$$

Where the parameters that belong to the transition probability densities are estimated separately in a first stage:

$$L_1(\theta) = \prod_{i=1}^{N_i} \prod_{t=1}^{T_i} f(s_{t+1}^i|s_t^i,\theta)$$

and then used in partial likelihoods of the first stage:

$$L_2(\theta) = \prod_{i=1}^{N_i} \prod_{t=1}^{T_i} p(d_t^i|s_t^i,\theta)$$

# 5 Estimation

Estimation of the cost parameters that maximize the likelihood of the observed sequence of well operating choices over the sample period. Well level heterogeneity is accounted for only to the extent that the dynamic programming model is estimated separately for different areas and types of wells. The wells are divided into groups depending on, (1) whether the well is an oil or gas well, (2) whether the well is on a pool with more than one well, (3) the royalty regime applicable (that is, whether the well was is in a pool discovered before or after a specific date), (4) the PSAC area that the well was drilled, and within these groups, (4) two clusters based on time invariants (for gas): compressibility, density and depth of well. Within these final clusters all wells are treated as homogeneous.

The operator's decision process is characterized by a parameter vector $\theta$ that consists of the first stage parameters of the price process $\theta_{P0}$, $\theta_{P1}$, and $\theta_{P\sigma}$, of the production density $\theta_{q0}$, $\theta_{q1}$, and $\theta_{q\sigma}$, and of the density of recoverable reserves $\theta_{Q0}$, $\theta_{Q1}$, $\theta_{Q2}$, and the second stage parameters of the profit function ($C$, $M_1$, $M_2$, $SC_{(1\to2)}$, $SC_{(2\to1)}$, $SC_{(1\to3)}$, $SC_{(2\to3)}$) as well as those parameters that are not estimated $\beta$, $b$.

## 5.1 First Stage Estimates

The first stage involves estimating the operator's beliefs about future prices and recoverable reserves. These beliefs are unobservable and subjective, but here we assume that the operator's beliefs are recoverable from objective probability measures estimated from the data ("rational expectations"). In estimating the first stage transition probabilities there is no need to solve for the fixed point of the Bellman equation. This first stage is indeed 3 stages of estimation for the transition probabilities of 2 of the state variables (price and remaining reserves). The state variable age in years is aggregated into different age groups ($1 \leq age < 5$, $5 \leq age < 15$, $15 \leq age < 30$ and $age \geq 30$). The transition probability of entering the next age group is $1/n_{years}$ where $n_{years}$ is the number of years in the current age group.

### 5.1.1 Transition in Remaining Reserves

To determine the transition probability of reserve size, first the probability of extracting quantity $q$ is estimated, followed by estimating the probability of reserve growth or revisions. The probability of extracting $q$, given reserves $\overline{Q}$, is estimated by:

$$\log q = \theta_{q0} + \theta_{q1} \log \overline{Q} + \theta_{q\sigma} \varepsilon \tag{5}$$

separately for each well group (based on oil or gas wells, whether there is more than one well in the pool, royalty regime, PSAC area and cluster), as well as each age group. With $\varepsilon$ as an identically and independently distributed $N(0,1)$ error. If there are less than 100 observations of production within a given cluster's age group, then observations from the age group of the unclustered well group is used.

The data on reserve estimates contain (1) estimates of the physical reserve, that is the actual oil or gas in place, and (2) the recoverable reserve, the portion of the physical reserve that is extractable. The estimates of the physical reserves increase or decrease because of production, new discoveries or a revision of the old estimate. The recoverable reserves change with the same factors but also with changes in technology and prices. In the estimation, the reserve estimates are converted to per well reserves, $\overline{Q}$, where the reserve estimates are divided by the number of wells in the pool. Therefore, as by construct of $\overline{Q}$, when a new well is drilled the reserves automatically decrease. For the time being, this decrease is estimated together with the decrease by reassessment of reserves, so that there is single a probability density function for the changes in reserves by reassessment, another well drilled in the pool, or both. To measure this transition probability density function (of discoveries, reassessments, technology or new wells drilled), the observations on per well initial established reserves $\overline{Q}_{IER}$ are used and not the observations on remaining reserves.

The vast majority of the time (see table 4) there is no change in per well initial established reserves, that is $\overline{Q}'_{IER}/\overline{Q}_{IER} = 1$. When there is an increase in reserves, $\log(\overline{Q}'_{IER}/\overline{Q}_{IER})$ can be approximated by an exponential distribution with rate parameter $\lambda_U$, different from when there is a decrease in reserves, and $-\log(\overline{Q}'_{IER}/\overline{Q}_{IER})$ can be approximated by a different exponential distribution with rate parameter $\lambda_L$. The rate parameters depend on the price of the hydrocarbon so that $f_{\overline{Q}}(\overline{Q}'|\overline{Q}, P)$ is estimated with an exponential regression:

$$f_{\overline{Q}}\left(\Delta|\overline{Q}, P\right) = \theta_{\overline{Q}} \exp\left(-\theta_{\overline{Q}}\Delta\right)$$

where $\Delta = \log(\overline{Q}'_{IER}) - \log(\overline{Q}_{IER})$ when estimating upwards revisions in reserves, and because the exponential density occurs in the positive domain, $\Delta = -(\log(\overline{Q}'_{IER}) - \log(\overline{Q}_{IER}))$ when estimating downward revisions.

$$\theta_{\overline{Q}} = \begin{cases} \theta_{\overline{Q}0} + \theta_{Q1}/P & \text{when } \log(\overline{Q}'_{IER}) - \log(\overline{Q}_{IER}) < 0 \\ \theta_{\overline{Q}0} + \theta_{Q1}P & \text{when } \log(\overline{Q}'_{IER}) - \log(\overline{Q}_{IER}) > 0 \end{cases}$$

Table 4: Parameters for Transition of Initial Established Reserves

| | Decrease $\left(\log(\overline{Q}'_{IER}/\overline{Q}_{IER}) < 0\right)$ | Increase $\left(\log(\overline{Q}'_{IER}/\overline{Q}_{IER}) > 0\right)$ | Stays the Same $\left(\log(\overline{Q}'_{IER}/\overline{Q}_{IER}) = 0\right)$ |
|---|---|---|---|
| Probability | 0.1101046 | 0.10421477 | 0.78568063 |
| $\theta_{Q0}$ | 0.90621261 | 0.27379018 | |
| $\theta_{Q1}^{\ddagger}$ | 0.00083298835 | 680.15871 | |
| $LL$ | -8426.5018 | -971.53257 | |
| $No.Obs.$ | 9347 | 8847 | 66698 |

‡ For decreases in reserves, $1/P$ is the regressor, whereas for increases in reserves, $P$ is the regressor (where $P$ is measured in millions of dollars per thousand m$^3$ of gas). The probability of increase, decrease or staying the same is the average occurrence observed in the data.

Reserve changes due to extraction are such that when the well is active, the quantity extracted is modeled as a random draw from a distribution that depends on the per well remaining reserves, $\overline{Q}$. It implies that the operator only chooses whether to extract and does not have control over quantity extracted. To defend this assumption, Table 5 shows that variables that the operator has no control over (such as depth, density, initial pressure, and compressibility) are very important in determining extraction. Price on the other hand, gives the surprising result that there is less extraction with higher prices in three of the specifications, and 1 percentage change in price leads to only .9-.19 percentage change in extraction.

However there is the problem with modeling extraction as only dependent on reserves, in that as described in the previous paragraph, the reserves are estimated reserves and prone to revisions. This implies that the quantity extracted depends on what is believed to be in place, not what is actually in place (this would mean that if estimates of reserves are revised to half that of what was believed, then extraction would decrease correspondingly). But to defend this choice for the model, many of the reserve estimates are changes due to revisions in the recovery factor, and therefore a change in recovery factor should change the quantity recovered.

#### 5.1.2 Transition in Price

Price follows an exogenous first order Markov process. It is a reasonable assumption for formation of expectations for future prices. The operator forms their expectation based on current prices. We model price according to a lognormal $AR(1)$ process:

$$\log P' = \theta_{P0} + \theta_{P1} \log P + \theta_{P\sigma} \varepsilon \tag{6}$$

where $\varepsilon$ is an IID $N(0,1)$. It is also truncated to the interval $[\underline{P}, \overline{P}]$.

Table 5: Log Log Maximum Likelihood Estimates for Quantity Extracted

| Oil Production | I | II | Gas Production | III | IV |
|---|---|---|---|---|---|
| Intercept | -0.25** | -7.15** | Intercept | -1.82** | 1.03** |
| | (0.057) | (1.09) | | (0.17) | (0.21) |
| $\overline{Q}_o$ | 0.38** | 0.31** | $\overline{Q}_g$ | 0.57** | 0.34** |
| | (0.0039) | (0.0046) | | (0.0023) | (0.0034) |
| $\text{Price}_o$ | -0.13** | -0.16** | $\text{Price}_g$ | 0.099** | -0.19** |
| | (0.015) | (0.017) | | (0.014) | (0.014) |
| $\text{Age}_o$ | -0.13** | -0.21** | $\text{Age}_g$ | -0.19** | -0.16** |
| | (0.0048) | (0.0058) | | (0.0045) | (0.005) |
| No. Of Wells in $\text{Pool}_o$ | | 0.072** | No. Of Wells in $\text{Pool}_g$ | | 0.013** |
| | | (0.0046) | | | (0.0056) |
| $\text{Depth}_o$ | | -0.40** | $\text{Depth}_g$ | | 0.15** |
| | | (0.047) | | | (0.022) |
| $\text{Porosity}_o$ | | -0.08** | $\text{Porosity}_g$ | | -0.20** |
| | | (0.018) | | | (0.018) |
| $\text{Density}_o$ | . | 0.65** | $\text{Density}_g$ | | 1.96** |
| | | (0.15) | | | (0.073) |
| Initial $\text{Pressure}_o$ | | 0.50** | Initial $\text{Pressure}_g$ | | 0.14** |
| | | (0.025) | | | (0.016) |
| No. Of Firms in $\text{Pool}_o$ | | -0.076** | No. Of Firms in $\text{Pool}_g$ | | -0.20** |
| | | (0.0096) | | | (0.013) |
| Wells per Firm | | 0.019** | Wells per Firm | | 0.0045 |
| | | (0.0027) | | | (0.0029) |
| $\text{Temperature}_o$ | | 0.058 | | | |
| | | (0.036) | | | |
| Water $\text{Saturation}_o$ | | -0.29** | | | |
| | | (0.015) | | | |
| | | | $\text{Compressibility}_g$ | | 4.47** |
| | | | | | (0.1) |
| No. of Observations | 84366 | 70180 | No. of Observations | 112375 | 89748 |
| Log Likelihood | -143000 | -117385 | Log Likelihood | -213786 | -166723 |
| R Square | 0.1365 | 0.1522 | R Square | 0.3689 | 0.2372 |

Dependent variable: natural logarithm of annual quantity produced. All variables are log transformed. ** parameter estimates are significantly different from zero at the 2.5% level. Standard errors are in parentheses.

Table 6: First Stage Estimates of Price Process

| | Oil Price | | Gas Price | |
|---|---|---|---|---|
| $\theta_{P0}$ | -1.6993 | (1.0039) | -0.9566 | (0.5733) |
| $\theta_{P1}$ | 0.8340 | (0.0955) | 0.8904 | (0.0616) |
| $\theta_{P\sigma}$ | 0.2014 | (0.0237) | 0.2167 | (0.0255) |
| $LL$ | -6.6146 | | -3.9766 | |
| $No.Obs.$ | 36 | | 36 | |

Estimates correspond to equation (6). Natural logarithm of gas price in millions of dollars per thousand$m^3$, and oil price in millions of dollars per barrel. Standard errors are in parentheses.

$$F_P(P'|P) = \frac{\Phi\left((\log(P') - \mu)/\sigma\right) - \Phi((\log(\underline{P}) - \mu)/\sigma)}{\Phi((\log(\overline{P}) - \mu)/\sigma) - \Phi((\log(\underline{P}) - \mu)/\sigma)}$$

where $\mu = a_0 + a_1 \log(P)$.

## 5.2 Second Stage

Here, the structural parameters $\theta = (C,\ M_1,\ M_2,\ SC_{(1\to 2)},\ SC_{(2\to 1)},\ SC_{(1\to 3)},\ SC_{(2\to 3)})$, are estimated by maximizing the likelihood that choice for well $(i)$ at time $(t)$ is decision $d$:

$$L(\theta) = \prod_{i=1}^{N_i} \prod_{t=1}^{T_i} p(d_t^i | P_t, Q_t^i, \theta)$$

where $p$ is the multinomial logit given in equation 4. For each iteration of the likelihood there is a nested subroutine to find the fixed point to the Bellman equation 3. The dynamic programming model is in discrete time and the operator decides the operating mode on a yearly basis. In reality this decision is in continuous time however, the classification by the Alberta Energy Resources Conservation Board is that an inactive well is one that has not reported any volumetric activity within the last 12 months (including production, injection or disposal). The action and states were defined as follows: for a given well in 1996, the decision is be the operating state observed in 1996, under the average wellhead price of oil in 1996, the quantity of remaining reserves in 1996, the age of the well in 1996, the operating state in 1995.

The royalty rate was calculated using the formulas specified by the Alberta Department of Energy[Fiscal, 2006]. The rates can range from 5% to 35%. The royalty rate is sensitive to the market price, age of the well and the volume of hydrocarbons produced. The royalty rate used in the model is different for different well ages, and sensitive to the price of hydrocarbon as well as production, however it is an approximation for the following reasons. As this model is based on the expected production, and not the actual production, the royalty rate is the expected royalty rate. Less of a problem is that there is a select price under which "old", "new", and "third tier" rates depend. This price is issued by the government and is slightly different every year, however for all years, the select price of 2006-2007 is used, as the past wellhead prices are inflated to 2007 prices anyhow. And finally, the model does not include royalty adjustments for deep marginal gas wells, $CO_2$ enhanced production, exploration wells, nor the reactivated well royalty exemption. This last omission of a program on reactivation might cause concern. The program exempts payment on royalties for the first 50318 barrels of oil from wells that have been reactivated after 24 months of inactivity. As this is a fixed amount of royalty relief for large and small wells, it is likely to be absorbed by making switching cost to reactivate these wells lower. There are three different royalty regimes for oil depending on when the well was drilled: "old oil" for wells drilled on pools discovered before 1974, "new oil" for wells drilled on pools discovered between 1974 and 1992, and “third tier oil” for wells drilled on pools discovered after 1992. While for gas there are: “old gas” for wells drilled on pools discovered after 1974 and ”new gas” for wells that were drilled on pools discovered after 1974. Therefore, the data is divided into samples according to the royalty regime that the wells face. The royalty rate is the same for old and new oil wells, only differing in the price range that the royalty rate becomes price sensitive. Below a certain quantity extracted, they face a flat rate, but once extraction exceeds a certain quantity, then the royalty depends on how much is extracted. Third tier oil differs from old and new oil in the quantity that the royalty rate stops being flat, and starts depending on quantity produced. The Alberta corporate income tax rate is 10% of taxable income while the federal corporate income tax rate is 22.12%. So the combined federal and provincial tax rate on corporate income is set at 32.12%[Corporate, 2007].

Estimating the discount factor, $\beta$, along with the switching costs is difficult: both a high reactivation cost and a low discount factor will prolong reactivation. After estimating the model with different fixed discount factors it was found that a discount factor of .75 results in slightly higher likelihood values. A discount factor of .75 corresponds to an annualized discount rate of 28.77% ($\beta = \exp(-r)$). This is consistent with Farzin [1985] who estimated a before tax discount rate in the oil and gas industry to be 25.4%.

The model was estimated using a multitude of different specifications for the cost function. A parsimonious specification that leads to timely convergence and high likelihood values is a lifting cost of $C = \theta_1/\overline{Q}^{1.5} + \theta_2(1+\theta_3)^A + I_s\theta_4 P + I_l\theta_5 P$, where $I_s$ is an indicator for small pools, and $I_l$ is an indicator for large pools. The lifting cost depends on the price of the hydrocarbon following the Alberta Department of Energy estimate that operating costs change 33.3% for a 100% change in price[7]. The average estimates across PSAC areas of the structural cost parameters for groups of wells depending on whether they are in pools with only one well, in pools with more than one well and royalty regime applicable (“new gas” rate for wells in pools discovered before 1974 or “old gas” rate for wells in pools discovered 1974 or after) are presented in Table 7. The annual fixed cost of leaving the well inactive, $M_2$, and operating, $M_1$, are negative because the they are interpreted relative to the abandonment cost which for identification was fixed at \$75,000, $SC_{(1,2\to 3)} = .075$. By construct of the multinomial logit (equation 4), identifying all fixed costs of the model is not possible because both $\pi$ and $a\pi + b$ will return the same decision rule. That is, the location and scale of the profit function is unidentifiable, and therefore the profit is normalized by a prespecified location and scale. The location is normalized by fixing the abandonment cost at \$75,000. The scale is normalized by assuming that the error term $\epsilon_t$ has a Type I extreme value distribution

[7] http://www.energy.gov.ab.ca/Oil/pdfs/RISConvTechInvestorCompar.pdf

Table 7: Average Parameter Estimates Across PSAC Areas

| Parameter | Gas | | | |
|---|---|---|---|---|
| | Single-Well Pool | | Multi-Well Pool | |
| | Old | New | Old | New |
| $M_2$ | -0.8575(0.47194) | -1.2733(1.1473) | -0.88349(0.21576) | -0.99264(0.65859) |
| $M_1$ | -2.0686(0.94006) | -2.3711(1.7753) | -1.7775(0.51933) | -1.8561(0.96013) |
| $SC_{1\to2}$ | 0.63457(2.6206) | 1.9351(1.4176) | 1.8946(0.96417) | 1.8884(0.62847) |
| $SC_{2\to1}$ | 3.9642(2.5828) | 2.3744(1.1437) | 3.8961(0.94022) | 3.0819(0.6325) |
| $\theta_1$ | 0.033624(0.038288) | 0.025886(0.038037) | 0.026775(0.035613) | 0.0075236(0.019562) |
| $\theta_2$ | 3.4151e-005(5.5667e-005) | 2.2643e-005(4.1093e-005) | 3.5018e-005(5.4777e-005) | 6.3948e-006(1.3679e-005) |
| $\theta_3$ | 0.22214(0.32161) | 0.43558(0.44218) | 0.2964(0.33726) | 0.34632(0.2726) |
| $\theta_4$ | 0.22607(0.18495) | 0.27097(0.2251) | 0.45286(0.2922) | 0.51002(0.17358) |
| $\theta_5$ | 0.23196(0.2426) | 0.26117(0.34751) | 0.35561(0.24709) | 0.48965(0.25942) |

Notes:In parenthesis are the standard deviations of the estimates across the seven PSAC areas, not standard errors of estimates.

with standard deviation, $b$ equal to 1. The variation across PSAC areas in the costs is far greater for wells that are on their own pool compared to those on pools with more than one well.

To test the dynamic programming model's ability to fit the data, the choice probabilities from the estimated dynamic programming model $p(d|s,\hat{\theta})$, are compared to the observed (non-parametric) estimates of the conditional choice probability function $\hat{p}(d|s)$. The nonparametric estimate $\hat{p}$ is the sample histogram of choices made in the subsample with state $s$. Following Rust and Phelan [1997] and Rothwell and Rust [1997], by sample enumeration, the nonparametric estimate of the choice probability is computed as:

$$\begin{aligned}\hat{p}(d|S) &= \int_{s\in S}\hat{p}(d|s)\hat{F}(ds|S)\\ &= \frac{1}{N}\sum_{i=1}^{N} I\{d_i = d, s_i \in S\}\end{aligned}$$

Where $\hat{F}(ds|S)$ is the nonparametric estimate of the conditional probability distribution of $s$ given $S$, equal to the number of observations in cell $ds$ divided by the total number of observations in cell $S$. And the dynamic programming estimates conditional probability, for each area and group of wells is:

$$\begin{aligned}p(d|S,\hat{\theta}) &= \int_{s\in S} p(d|s,\hat{\theta})\hat{F}(ds|S)\\ &= \frac{1}{N}\sum_{i=1}^{N} p(d|s,\hat{\theta}) I\{s_i \in S\}\end{aligned}$$

Where $p(d|s,\hat{\theta})$ is the probability given by 4 and $\hat{\theta}$ are the estimates of the structural parameters. As different dynamic programming models were estimated for the different groups

Table 8: Actual versus Predicted Choice Probabilities: Full Sample

| **When Current State** | Observed | Expected |
|---|---|---|
| **Active or Inactive** | | |
| Pr(Active) | 0.6866676143 | 0.6856170908 |
| Pr(Inactivate) | 0.3038522099 | 0.304437222 |
| Pr(Abandon) | 0.009480175826 | 0.009945687183 |
| No. Obs. | | 182908 |
| $\chi^2$ | | 4.485313106 |
| $Pr(\chi^2(2)) \geq 4.49)$ | | 0.106176067 |
| **Active** | | |
| Pr(Active) | 0.9425631963 | 0.9399397287 |
| Pr(Inactivate) | 0.05291602234 | 0.05479487187 |
| Pr(Abandon) | 0.004520781393 | 0.005265399472 |
| No. Obs. | | 126748 |
| $\chi^2$ | | 22.44043412 |
| $Pr(\chi^2(2)) \geq 22.44)$ | | 1.340051981e-005 |
| **Inactive** | | |
| Pr(Activate) | 0.109134615 | 0.111634083 |
| Pr(Inactive) | 0.870192308 | 0.867857247 |
| Pr(Abandon) | 0.0206730769 | 0.0205086699 |
| No. Obs. | | 56160 |
| $\chi^2$ | | 3.56971746 |
| $Pr(\chi^2(2)) \geq 3.57)$ | | 0.167820767 |

Notes: $\chi^2$ was calculated as $N \sum_{i=1}^{3} (NP - DP)^2/DP$, where $N$ is the number of observations.

(based on location or type of well), the conditional choice probabilities differed by these groups. Table 8 compares the choice probability estimates in the case where $S$ is a collection of all possible $s$ cells, and as different dynamic programming models were estimated for different groups of wells, different conditional probabilities were calculated, and then averaged.

# 6 Simulation

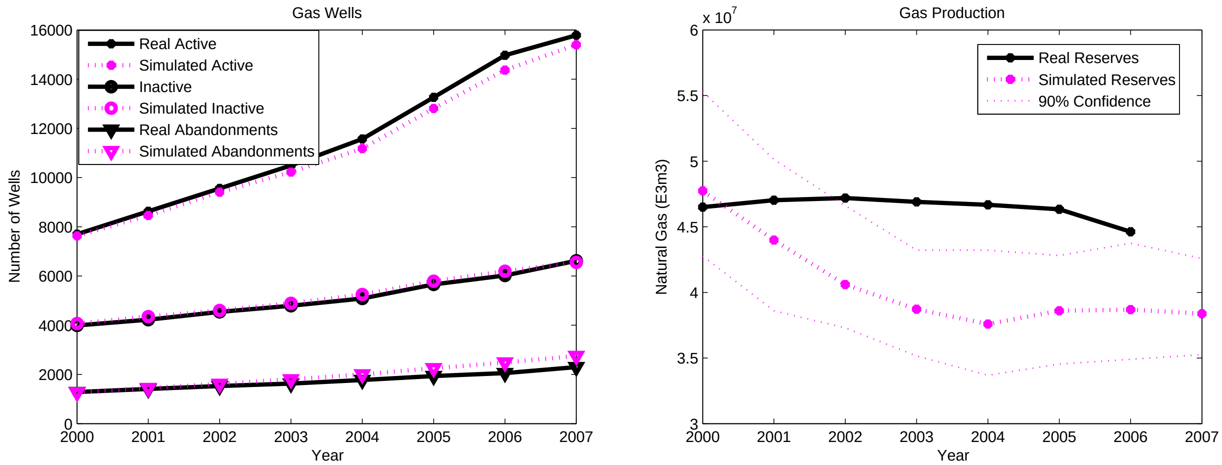
## 6.1 Re-Simulating the Data

Figure 2 illustrates the dynamic programming model's ability to forecast the Albertan conventional oil and natural gas industry (still to complete other PSAC areas, and oil wells). The simulation begins with the state of the industry in 2000 as a starting point. Each well enters the simulation with only information from the first year that it appears, that is, information from 1999, or if drilled post-1999, the year it was drilled (a well only appears in the dataset one year after it is drilled). From this first year the current operating state, age and remaining reserves then progress according to the dynamic programming model. The only state variable that is not simulated is the price of the hydrocarbon; each year of the simulation the real wellhead price is used. The estimated DP model returns the choice probabilities for operating state under all possible combinations of state variables. In the simulation the choice of operating state depends on where a $U(0,1)$ draw falls within intervals designated by the state dependent choice probabilities. No matter the choice taken, in each period the age of the well increases by a year, and a draw from $U(0,1)$ determines whether there are revisions to reserve estimates (and if so, a further draw from the exponential distribution determines by how much). If the operating choice is to extract, and the well is the only well in the pool, the remaining reserves decrease by a random draw from $f_q(\cdot|Q,A)$. If there is more than one well in the pool then the reserves decrease by this amount under any operating choice. In order to make the simulated comparable to the actual data, wells that did not appear in all years were dropped from the simulation (not the estimation); there being any intermittent missing years is due to missing information on some pools in some years. Contained in the simulation are separate estimation results for wells in single-well and multi-well pools, as well as clusters (if clustering increased the likelihood value).

The estimated parameters from observations from 2000 to 2007 are used to predict out of sample observations from 1993-1999. As seen in Figure 3 the model is able to predict out of sample observations.

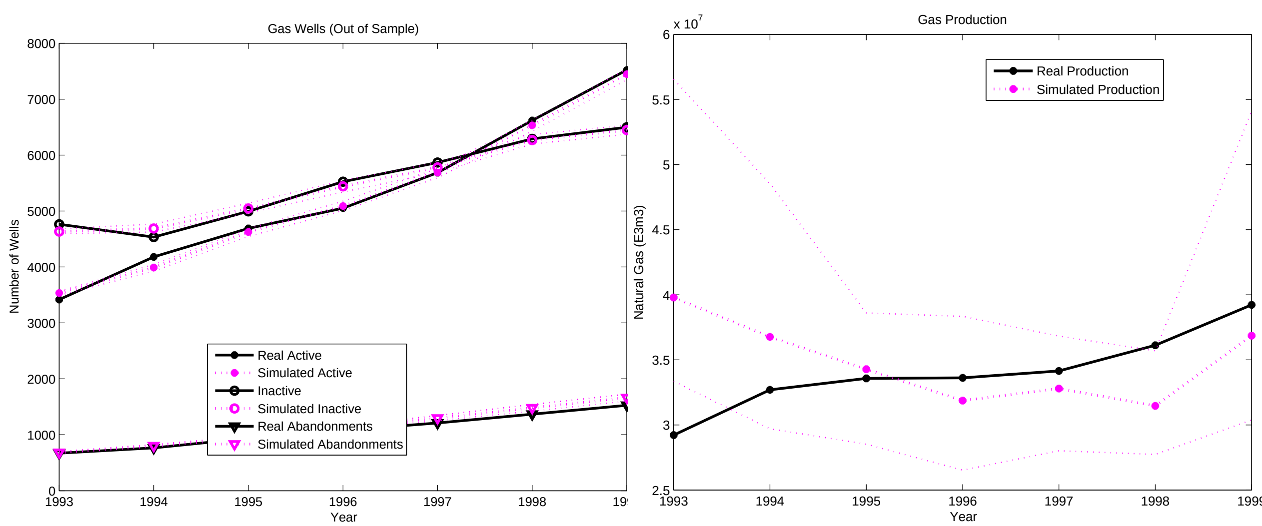
### 6.1.1 Counterfactual Simulations

Before implementing policies to deal with the accumulation of inactive wells (either policies to induce production or permanent closure), it is important to know whether exogenous changes in prices or technology would change the number of inactive wells. Two hypothetical scenarios are examined to see if the probability of inactivity for gas wells is significantly reduced. The first scenario (as seen in column 3 of Table 9) is the hypothetical situation of technology increasing the




(a) Composition of wells (b) Production

Figure 2: Real and Simulated Gas Wells and Production




(a) Composition of wells (b) Production

Figure 3: Real and Simulated: Out of Sample Prediction 1993-1999

Table 9: Actual versus Predicted Choice Probabilities Under Different Scenarios

| **Probability** | Observed | Expected | $\overline{Q} \times 100$ | $P_{2005} \times 4$ | No Taxes |
|---|---|---|---|---|---|
| Activate | 0.6867 | 0.6856 | 0.7711 | 0.7435 | 0.8034 |
| Inactivate | 0.3039 | 0.3046 | 0.2251 | 0.2495 | 0.1940 |
| Abandon | 0.0095 | 0.0098 | 0.0038 | 0.0070 | 0.0026 |
| No. Obs. | 182908 | | | | |

remaining established reserves for all observations by 100. Increasing the reserve size by a factor of 100 only reduces the probability of inactivity by 26%, indicating the extent of the hysteresis of inactivity. The second scenario examines the choice probabilities when the price of gas is four times the highest observed price (ie. the 2005 price, \$293.90$E^3m^3$, multiplied by 4). As seen from column 4 of Table 9, increasing the price by four only decreases the probability of inactivity by 18%.

The extent of the hysteresis is also illustrated by a policy simulation of no taxes (that is, the corporate income tax and the royalty rate were set to zero). Although reducing the tax rate to zero is unrealistic, it is illustrative to see how removing a tax of 32.12% and royalty of 5% to 35% affects the proportion of inactive wells. As shown in column 5 of Table 9, the expected probability of inactivity is decreased by 36%. Future research will involve evaluating a tax schedule on inactive wells.

# 7 Conclusion

A dynamic discrete choice model for the operating state of oil and gas wells in Alberta is estimated in order to determine what conditions are necessary to decrease the number of inactive wells. The operating decisions taken for a subsample of wells in Alberta can be replicated by a model where well operators are dynamic optimizers with an annual discount rate of about 29%. Within-sample goodness of fit tests show that the model is able to closely predict actual operating choices. The model is further validated using an out-of-sample prediction by closely predicting decisions for data not used in the estimation of the parameters. Preliminary findings show that the probability of choosing to leave a well inactive does not decrease proportionately with an increase in price or an increase in recoverable reserves. In order to see a substantial reduction in the number of inactive wells a dramatic improvement in extraction technology and wellhead price would be needed. Should such an extreme increase in recoverable reserves or prices not be foreseeable then there is an argument for implementing policies to reduce the number of inactive wells.

# References


J.R. Arrington. Predicting the size of crude reserves is key to evaluating exploration programs. *The Oil and Gas Journal*, 58(9):130–134, 1960.

D. Beliveau and R. Baker. Reserves growth: Enigma, expectation or fact? Technical report, Prepared for presentation at the Society of Petroleum Engineers Annual Technical Conference and Exhibition, 2003. URL `http://www.epiccs.com/resources/publications/SPE84144.pdf`.

J. Boyd. Financial responsibility for environmental obligations: Are bonding and assurance rules fulfilling their promise? Discussion Papers dp-01-42, Resources For the Future, 2001. URL `http://ideas.repec.org/p/rff/dpaper/dp-01-42.html`.

M. J. Brennan and E. S. Schwartz. Evaluating natural resource investments. *Journal of Business*, 58(2): 135–157, 1985.

J.M. Chermak and R.H. Patrick. A well-based cost function and the economics of exhaustible resources: The case of natural gas. *Journal of Environmental Economics and Management*, 28:174–189, 1995.

Corporate. Royalty information briefing 6-corporate income tax, alberta royalty review. Technical report, Alberta Department of Energy, 2007. URL `http://www.assembly.ab.ca/lao/library/egovdocs/2007/aleo/159246.pdf`.

A. K. Dixit and R. S. Pindyck. *Investment Under Uncertainty.* Princeton University Press, Princeton, NJ, 1994.

EIA. World dry natural gas production, most recent annual estimates, 1980-2007, 2008. URL `http://www.eia.doe.gov/emeu/international/gasproduction.html`.

EIA. Petroleum (oil) production, international petroleum monthly, January 2009.

Alberta Department of Energy. Alberta royalty review 2007: Royalty information series. Technical report, 2007.

ERCB. Alberta's energy reserves 2007 and supply/demand outlook 2008-2017. Technical report, Energy Resources Conservation Board, 2008.

ERCB013. Suspension requirements for wells. Technical report, Alberta Energy Resources Conservation Board, 2004. URL `http://www.ercb.ca/docs/documents/bulletins/bulletin-2004-29.pdf`.

Y. Hoosein Farzin. *Competition and Substitutes in the Market for an Exhaustible Resource.* Jai Press Inc., Greenwich, CT, 1985.

D.F. Ferreira, S. B. Suslick, and P.C.S.S. Moura. Analysis of environmental bonding system for oil and gas projects. *Natural Resources Research*, 2003.

Fiscal. Oil and gas fiscal regimes: Western canadian provinces and territories. Technical report, Alberta Department of Energy, 2006. URL `http://www.energy.gov.ab.ca/Tenure/pdfs/FISREG.pdf`.

M. Foss and D.V. Gordon. The cost of lifting natural gas in alberta: A well level study. 2007.

A. Gamba and M. Tesser. Structural estimation of real options models. August 2007.

R. Kubichek, J. Cupal, W. Iverson, S. Choi, and M. Morris. *Identifying Ground-Water Threats from Improperly Abandoned Boreholes.* 1997.

John R. Livernois. Empirical evidence on the characteristics of extractive technologies: The case of oil. *Journal of Environmental Economics and Management*, 14(1):72–86, March 1987.

J.R. Livernois and R.S. Uhler. Extraction costs and the economics of nonrenewable resources. *The Journal of Political Economy*, 95(1):195–203, 1987.

C. F. Mason. Nonrenewable resources with switching costs. *Journal of Environmental Economics and Management*, 42:65–81, 2001.

MineralAct. Mines and minerals act: Petroleum royalty regulation. Technical report, ALBERTA REGULATION 248/90, 2008. URL http://www.qp.gov.ab.ca/documents/Regs/1990$_2 48.cfm?frm_i sbn =$ 9780779729098.

A. Moel and P. Tufano. When are real options exercised? an empirical study of mine closings. *The Review of Financial Studies*, 15(1):35–64, 2002.

D. Morehouse. The intricate puzzle of oil and gas "reserves growth". *Energy Information Administration, Natural Gas Monthly*, pages vii–xx, 1997.

Orphan. Orphan well association 2005/06 annual report, 2006. URL http://www.orphanwell.ca/2005-06%20OWA%20Ann%20Rpt%20Final.pdf.

J. L. Paddock, D. R. Siegel, and J. L. Smith. Option valuation of claimes on real assets: The case of offshore petroleum leases. *Quarterly Journal of Economics*, 103(3):479–508, 1988.

V. Parente, D. Ferreira, E. Moutinho dos Santos, and E. Luczynski. Offshore decommissioning issues: Deductibility and transferability. *Energy Policy*, 34:1992–2001, 2006.

Geoffrey Rothwell and John Rust. On the optimal lifetime of nuclear power plants. *Journal of Business and Economic Statistics*, 15(2):195–208, 1997.

RRC. Railroad commission plugs record number of abandoned wells in fiscal year 2006, 2006. URL http://www.rrc.state.tx.us/news-releases/2006/100606.html.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 55(5):999–1033, 1987.

John Rust. Maximum likelihood estimation of discrete control processes. *SIAM Journal on Control and Optimization*, 26(5):1006–1024, 1988.

John Rust. Structural estimation of markov decision processes. In R. F. Engle and D. McFadden, editors, *Handbook of Econometrics*, volume 4, chapter 51, pages 3081–3143. 1994.

John Rust and Christopher Phelan. How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica*, 65(4):781–831, 1997.

M. E. Slade. Valuing managerial flexibility: An application of real-option theory to mining investments. *Journal of Environmental Economics and Management*, 41:193–233, 2001.

M. Verma and M.E. Henry. Historical and potential reserve growth in oil and gas pools in saskatchewan. Miscelaneous report, Saskatchewan Geological Survey, 2004.

M.L Williams, C.R. Matthews, and T Garza. Well plugging primer. Technical report, Railroad Commission of Texas, Well Plugging Section, Oil and Gas Division, 2000. URL http://www.rrc.state.tx.us/divisions/og/key-programs/plugprimer1.pdf.