CAUSAL IMPACTS OF FINANCIAL MARKET REGULATION: EVIDENCE FROM CENTRAL CLEARING AND METHODOLOGICAL ADVANCES IN DIFFERENCE-IN-DIFFERENCES

By

Arnob L. Alam

Submitted to the College of Arts and Sciences

of the American University

in Partial Fulfillment of

the Requirement for the Degree of

Doctor of Philosophy

In

Economics

Chair: ____________________________________

Professor Amos Golan

____

Professor Robin Lumsdaine

____

Professor Kara Reynolds

____

Dean of College

____

Date

2025

American University

Washington, D.C. 20016

DEDICATION

CAUSAL IMPACTS OF FINANCIAL MARKET REGULATION: EVIDENCE FROM CENTRAL CLEARING AND METHODOLOGICAL ADVANCES IN DIFFERENCE-IN-DIFFERENCES

BY

Arnob L. Alam

ABSTRACT

This dissertation consists of two essays exploring the causal impacts of financial market regulations using causal inference methods. The first essay investigates the impact of mandatory central clearing on prices, volatility, and liquidity in the interest rate swaps market. Utilizing a classical difference-in-differences framework, it leverages regulatory variation across currency denominations to isolate causal effects of the clearing mandate. Empirical findings indicate an increase in interest rate swap premia with the adoption of central clearing, consistent with the theoretical model. Market liquidity and volatility does not show any improvement with the adoption of clearing, contrary to theoretical models.

The second essay examines the assumptions underlying the classical difference-in-differences method employed in the first essay. It specifically addresses the parallel trends assumption, proposing methodological innovations that relax this restrictive assumption. Specifically, it proposes using generalized maximum entropy (GME) based approach to estimating selection probability into treatment. Then, using a weighting scheme proposed by Abadie (2005), the average treatment effect on the treated can be estimated. If heterogeneous treatment effects are of concern, the GME-based method can be extended to estimate the conditional average treatment effect. Using simulations, this essay demonstrates that conventional difference-in-differences estimates are biased when the parallel trends assumption is violated. It shows that GME-based methods outperform in recovering the correct model parameters, especially for small data sets or highly collinear covariates. The method is applied to a real-world dataset where an estimate of the true treatment effect is available from a randomized control trial. The GME-based method can recover treatment effects closer to the best estimate than classical difference-in-differences.

Together, these essays enhance our understanding of financial market regulation effects and advance econometric methodologies, enabling more accurate causal inference in economic research.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

ABSTRACT .... *ii*

ACKNOWLEDGEMENTS .... *iii*

TABLE OF CONTENTS .... *iv*

LIST OF TABLES .... *vii*

LIST OF ILLUSTRATIONS .... *ix*

Essay 1: The Impact of Central Clearing on the Interest Rate Swaps Market *1*

Introduction *1*

1 Background *1*

1.1 Interest Rate Swaps **1**

1.2 Central Clearing **3**

1.3 Regulatory Background **7**

1.3.1 US Context 7

1.3.2 International Context 10

1.4 Review of Literature **12**

1.4.1 Interest Rate Swaps 12

1.4.2 Liquidity 13

1.4.3 Price Volatility 13

1.4.4 Systemic Risk and Contagion 14

1.4.5 Central Clearing 15

2 Theory *16*

2.1 Model of Swaps Pricing **16**

2.1.1 Pricing Without Credit Risks 16

2.1.2 Pricing with Counterparty Risk (Credit/Debit Valuation Adjustment) 17

2.1.3 Pricing Under Central Clearing 18

2.2 Model of Liquidity (Bid-Ask Spreads) **19**

2.2.1 Liquidity with no Counterparty Risk 19

2.2.2 Sensitivity to Specification of Utility Function 22

2.2.3 Liquidity with Counterparty Risk 24

2.3 Model of Price Volatility 25

2.3.1 Volatility without Counterparty Risk 25

2.3.2 Price Volatility in Markets with Counterparty Risk 26

2.4 Summary of Theoretical Findings 26

3 Data 27

3.1 Data Sources and Collection 27

3.2 Raw Data Description 28

3.3 Data Filtering and Cleaning 34

3.4 Yield Curve Construction 34

3.5 Descriptive Statistics 38

3.6 Limitations 40

4 Identification Strategy 40

4.1 Pricing 41

4.2 Liquidity 49

4.3 Volatility 55

5 Results 58

5.1 Pricing 58

5.2 Liquidity 66

5.3 Volatility 74

6 Conclusion 78

Essay 2: An Information Theoretic Extension to Conditional Differences-in-Differences 79

1 Review of Literature 81

2 Theory 82

3 Simulation *87*

3.1 First Stage 90

3.2 Second Stage 91

3.3 Extension: Sub-population treatment effects 92

4 Application to Real World Data *95*

5 VI. Conclusion *98*

Appendix A Contract Characteristics *100*

Definitions *101*

Appendix B Robustness Tests *104*

REFERENCES *112*

LIST OF TABLES

Table 1 Major Rule-Making Areas of the Dodd-Frank Act. 9

Table 2 Summary of Central Clearing Requirements in Major Financial Centers. 11

Table 3 Number of contracts and notional values by clearing status and reference rate for USD and CAD IR swaps (unfiltered data set). 30

Table 4 Sample data for construction of USD and CAD swaps curves. 36

Table 5 Selected contarct characteristics. 39

Table 6 Parallel Trends Pre-Trend tests. 44

Table 7 Placebo diff-in-diff results. 47

Table 8 diff-in-diff results for prices pooling all phases together. 60

Table 9 diff-in-doff results for prices by each implementation phase separately. 62

Table 10 diff-in-diff using alternative Currency Pair 64

Table 11 diff-in-diff analysis of relative bid-ask spreads. 68

Table 12 diff-in-diff analysis of Roll's Measure. 70

Table 13 diff-in-diff analysis of Amihud's Illiquidity Measure. 72

Table 14 diff-in-diff analysis of volatility of the fixed rate for USD and CAD IR swaps contracts 75

Table 15 diff-in-diff analysis of price volatility during second Grexit referendum 77

Table 16 Baseline simulation settings. 90

Table 17 Modifications to the baseline case for specific scenarios. 90

Table 18 Model Performance for Stage 1 over 5,000 Monte Carlo simulation runs. 91

Table 19 Second Stage regression results. 92

Table 20 Estimate of the conditional average treatment effect. 95

Table 21 Replication of LaLonde (1986) results 96

Table 22 Treatment effects using Abadie's weighting scheme (with GME first step). 98

Table 23 Contract characteristics for common contracts (Source: Bloomberg L.P.) 100

Table 24 Diff-in-diff results after including all pricing outliers 104

Table 25 Difference-in-differnces results after including pricing outliers. By phase analysis. 106

Table 26 Diff-in-diff results excluding the first day the treatment went into effect 108

Table 27 Diff-in-diff with alternative model specifications 110

LIST OF ILLUSTRATIONS

Figure 1 Example of obligations between three firms 6

Figure 2 A dealer-intermediated market witouth central clearing. 17

Figure 3 Dealer-based Market with Central Clearing. 19

Figure 4 Pre-trends for swap pricing for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation. 46

Figure 5 Pre-trends for Relative bid-ask spreads for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation. 52

Figure 6 Pre-trends for Roll's Measure for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation. 53

Figure 7 Pre-trends for Amihud Illiquidty Measure for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation. 54

Figure 8 Pre-trends for Realized Volatility for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation. 57

Essay 1: The Impact of Central Clearing on the Interest Rate Swaps Market

ARNOB L. ALAM

Introduction

The 2006-2008 financial crisis, the most severe economic downturn since the Great Depression, led to the passage of the Dodd-Frank Wall Street Reform and Consumer Protection Act (DFA). A key provision of the DFA required certain financial contracts to be cleared through a central counterparty (CCP). This study investigates the causal impact of this clearing mandate on prices, volatility, and liquidity in the interest rate (IR) swaps market, a major derivatives market used for hedging or speculating on interest rate risk. Despite extensive theoretical literature on central clearing, empirical studies are limited. Earlier research focused on the credit default swaps (CDS) market using event studies. Event studies cannot isolate causal impacts due to potential confounding factors. This essay addresses the gap in the literature by (1) examining the IR swaps market, which is larger and more widely used than the CDS market and (2) using a diff-in-diff approach to identify causal effects of the clearing mandate. Leveraging the fact that initial central clearing rules targeted IR swaps in the four largest currencies traded in the US but did not apply to contracts denominated in other currencies, this essay plausibly identifies the causal impact of the regulation on pricing, liquidity, and price volatility in the IR swaps market using a difference-in-differences approach.

The essay is organized as follows: section 1 provides background on the IR swaps market, the financial crisis, and the clearing mandate's role in post-crisis market reforms; section 2 develops the theory of pricing, price volatility and liquidity for IR swaps under a clearing mandate; section 3 discusses the identification strategy; section 3 details my data; section 5 discusses the results and section 6 concludes.

1 Background

1.1 Interest Rate Swaps

IR swaps are financial derivatives used to hedge or speculate on interest rate movements. The three most common types of IR swaps include "vanilla" fixed-for-floating swaps, basis swaps, and cross-currency basis swaps, with vanilla fixed-for-floating swaps being the most prevalent ("OTC Derivatives Statistics at End-June 2024" 2024). In this type of swap, one party exchanges fixed-rate coupon payments for floating-rate payments on a notional principal (Skarr and Szakaly-Moore 2007)[1]. Firms can use these instruments to convert floating-rate risk to fixed-rate risk, and vice versa. As a concrete example, imagine firm A can borrow at the London Interbank Offer Rate (LIBOR, a common variable interest rate used by banks when lending money to each other) or a fixed rate of 2.0%, while firm B can borrow at LIBOR + 0.25% or a fixed rate of 1.75%. Suppose firm A prefers borrowing at a fixed-rate and firm B prefers borrowing at a floating-rate (this could be because firm A owns fixed-income securities while firm B owns assets that pay a variable rate, and the firms would like to match their assets with their liabilities). Despite their preferences, firm A has a comparative advantage in borrowing at a floating-rate, and firm B in borrowing at a fixed-rate. To achieve their preferred arrangements, the firms can enter into an IR swap agreement with a $1M notional principal, where firm A receives a floating rate of LIBOR from firm B and pays a fixed rate of 1.75% to firm B. This transforms firm A's floating-rate liability into a fixed-rate one and vice versa for firm B. The IR swaps market allows firms to borrow in the market they have a comparative advantage in and trade for their preferred interest rate arrangement.

IR swaps can be bespoke contracts, customizable to individual economic needs (Loon and Zhong 2016). As the largest over-the-counter (OTC) swaps market, it accounted for $465 trillion of the $601 trillion global OTC swaps market in 2010 (von Kleist and Mallo 2011) (the IR swaps market had increased to $715 trillion by June 2023 according to an updated version of the same report). For many currencies, there are "standardized" contracts, which have common features and are the most heavily traded (Fett and Haynes 2017). During the period studied in this essay, the standard US Dollar (USD)-denominated IR swaps contract had semiannual payments for one leg and quarterly payments for the other leg, with the 3-month USD-LIBOR curve used both as the floating-rate reference and for discounting future cash flows (see section 2.1 for further explanation). The standard Canadian Dollar (CAD)-denominated contract used 3-month Canadian Dollar Offer Rate (CDOR) as the reference floating rate. In addition to the currency, reference rates and payment frequency, there are many other contract details (such as day-count conventions, settlement and termination rules) that need to be specified, and these are listed in more detail in

[1] The principal is "notional" because unlike a real bond it is never exchanged. It is only used to calculate fixed and floating rat payments.

Appendix A. The CAD- and USD-denominated standard contracts use the ISDA Master Agreement, which details these contract specifications (Minton 1997). Although contract specifications can be customized to meet the requirements of the counterparties, such non-standard contracts are likely to be less liquid than the standard contracts. Standard contracts denominated in other currencies (e.g. Euro [EUR], British Pound [GBP], and Japanese Yen [JPY]) have their own conventions as well (and these conventions are documented in Appendix A).

The IR swaps market is dealer-dominated, with dealer-customer and dealer-dealer trades accounting for 80% of notional value (Bolandnazar 2020). ~~Bolandnazar finds that~~ 50% of trades (by notional value) are executed by the largest seven dealers, indicating market concentration among a few dealers. This concentration can impact pricing and market stability in several ways. Larger dealers might be able to reduce search costs by easily finding a counterparty from their large client base. They could also reduce costs by economizing over administrative and warehousing costs of contracts. However, because of their market position, they might have market power and be able to charge a premium over the price that would prevail in competitive markets. The failure of a large dealer (or a dealer's major counterparty) could also drastically reduce liquidity in the system and increase transactions costs.

1.2 Central Clearing

When a swap is cleared, the initial contract between the two parties is replaced (novated) by two contracts between each party and a central clearinghouse/derivative clearing organization (CCP, DCO or clearinghouse) (Duffie and Zhu 2011; Duffie, Scheicher, and Vuillemey 2015). The clearinghouse becomes the counterparty for each leg (that is, receiving the fixed-rate payments from one party and paying the floating-rate payments to that party, while also receiving the floating-rate payments from the other party and paying it the fixed-rate payments). Under ordinary circumstances, the clearinghouse is a sort of "pass-through" organization that transmits payments from one counterparty to the other. However, if one party fails to meet their contractual obligation, the clearinghouse can still make sure the other party gets paid (Pirrong 2011). For this purpose, CCPs practice risk-control measures and have additional resources to make a counterparty whole in case of default[2]. When counterparties clear their trade through a clearinghouse, they must put

[2] The clearing counterparty is usually a dealer who is a clearing member at the CCP. "A clearing member is usually a trade intermediary that can deal directly with the CCP. Trade intermediaries that are not clearing members must clear their trades through a trade intermediary that is a

up collateral (initial margin) and contribute to a default fund. In case the risk position of the counterparty changes, it can be required to put up additional collateral (variation margin). The CCP also has default fund contributions from other members, its own equity (CCP capital), and access to other lines of credit (such as the Federal Reserve discount window). The combination of these resources makes it unlikely that the failure of one counterparty would drastically affect the whole market. Since clearing members can lose their contribution to the default fund in case of the failure of a counterparty, clearing mutualizes counterparty risk among the members of the CCP.

In addition to financial resources, CCPs exercise prudent risk-control measures such as monitoring members trading positions and liquidating distressed assets in an orderly fashion (Pirrong 2011). Since the CCP can observe all trades that it is clearing, it has a better picture of overall riskiness (compared to a bilateral market, where one party is generally unaware of other trades its partner is entering, and thus does not have a thorough understanding of its partner's riskiness).

Clearing can reduce demand for collateral through a practice called netting (Duffie, Scheicher, and Vuillemey 2015). There are two types of netting practices common in the industry: cross-product netting and multilateral netting. For a CCP that clears multiple types of contracts (e.g., IR swaps, forward rate agreements, overnight-index swaps, credit default swaps, etc.) cross-product netting involves netting across different derivatives products. For example, if firm A owes the CCP \$10 million in collateral for IR swaps, but the CCP owes firm A \$8 million for CD swaps, then firm A can just pay the CCP \$2 million in net collateral.

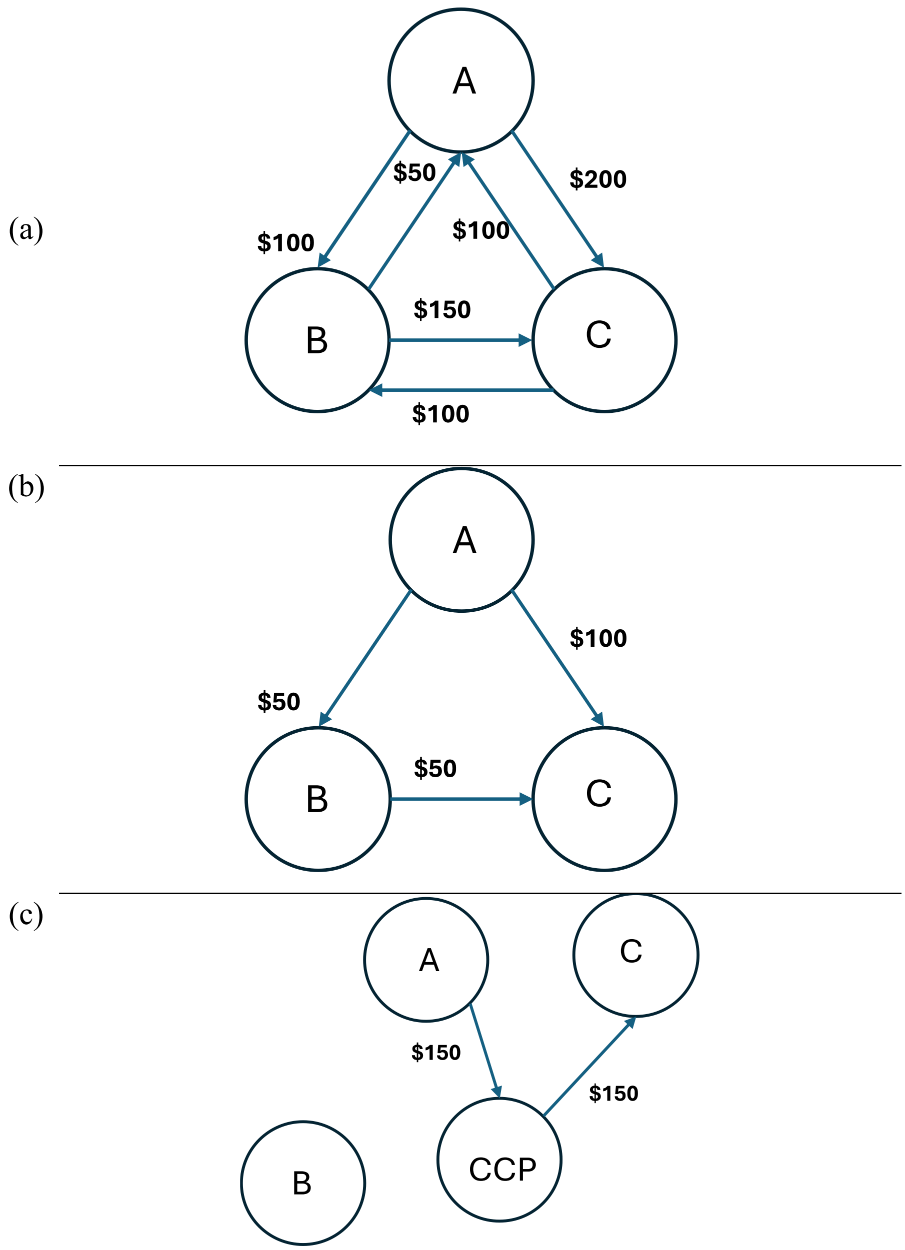
Multilateral netting involves netting payments across multiple firms. Consider the following example involving 3 firms. The set of obligations between the firms are as follows: firm A owes firm B \$100 million and firm C \$200 million; firm B owes firm A \$50 million and firm C \$150 million; firm C owes firm A \$100 million and firm B \$100 million. The total collateral demand in the system is \$700 million. This initial set of obligations is visualized in Figure 1a, where the arrows indicate the direction of the obligation (which firm owes who). Without multilateral netting, the firms can still engage in bilateral netting, as shown in Figure 1b. In a bilateral netting regime, the firms "subtract" or "net out" their payment to each counterparty. Thus, the following payments would be made: firm A would pay firm B (\$100 - \$50) = \$50 million and firm C (\$200 - \$100) =

clearing member" (McPartland 2009). Trade intermediaries that are clearing members will collect collateral from their non-clearing member clients and pass it on to the CCP.

$100 million; firm B would pay firm C ($150 - $100) = $50 million. The total collateral demand would be $200 million. As shown in the figure, under this arrangement, firm B acts like a pass-through entity that collects payment from firm A and transmits it to firm C. However, if firm B is unable to make the collateral payment, firm C loses some of the collateral it is due. Multilateral netting can eliminate this payment from firm B to firm C (with the CCP now acting as the pass-through entity). Under multilateral netting Firm A would pay the CCP $150 million and the CCP would pay firm C $150 million (while firm B would not make any payments at all). The total collateral demand would be $150 million. Figure 1c graphically depicts this multilateral netting scenario.

Figure 1 Example of obligations between three firms

(a) without any netting (b) with bilateral netting (c) with multilateral netting. Arrows indicate the direction of obligations (for example in (a) firm C owes firm B $100 million and firm B owes firm A $50 million). Under bilateral netting, firms "net out" the obligations with each counterparty on a bilateral basis. Thus, the set of two obligations between firm B and C is replaced by one obligation of $50 million from firm B to C, as shown in (b). Under central clearing, the payment from firm A to firm B can be eliminated and the clearinghiouse can simply collect $150 million in collateral from firm A and pass it directly to firm C, as shown in (c).



Originally created for members of futures and equities exchanges (Bernstein, Hughson, and Weidenmier 2019), clearinghouses became more significant with regulations like the DFA (2010) and European Market Infrastructure Regulation (EMIR, 2012) mandating central clearing of derivatives (Menkveld and Vuillemey 2021). Mandated clearing can have macro and micro effects on the swaps market (Pirrong 2011). At the macro level, clearing could reduce volatility but also strain the market through collateral demand during volatile or illiquid periods. Large enough losses could threaten clearinghouse solvency, transmitting effects to all members. At the micro level, central clearing may change the types of trades firms enter, potentially leading to riskier trades due to mutualized default risk (adverse selection) and riskier post-trade activities (moral hazard). Clearing is subject to economies of scale and scope, which could lead to natural monopolies. However, regulators are likely to prevent this through antitrust regulations and "local clearinghouse" requirements (that is, even though a single clearinghouse for both the US and Europe might have lower costs, US and EU regulators might require separate clearinghouses in each jurisdiction) (Benos et al. 2019). While clearinghouses can reduce default risk and collateral demand, they also require resources for risk management activities, which may increase trading costs.

1.3 Regulatory Background

1.3.1 US Context

Following the 2008 financial crisis, Congress passed the DFA to reform the entire US financial system. Since the OTC derivatives markets played a large role in the crisis, DFA aimed to significantly change how these markets worked. Key objectives included improving trade data availability for regulators and market participants, requiring real-time reporting of certain trade characteristics, and mandating confidential trade data reporting to swaps data repositories and regulators (*Dodd-Frank Wall Street Reform and Consumer Protection Act* 2010). To reduce default risk for large swaps dealers, DFA requires dealers to register with the Commodities Futures Trading Commission (CFTC) or the Securities and Exchange Commission (SEC), adhere to internal business conduct standards and maintain adequate capital (Commodity Futures Trading Commission, n.d.). To enhance liquidity, price discovery and transparency, it encourages trading to take place in centralized Swaps Execution Facilities (SEFs, usually electronic trading venues) or Designated Contract Markets (DCMs). To make trade data more readily available, it requires near real-time reporting and dissemination of price information to Swaps Data Repositories (SDRs) and submitting additional data (called primary economic terms) to SDRs and regulators in a timely fashion. Furthermore, the DFA mandates most contracts be centrally cleared (and for uncleared contracts, requires parties to post regulatory margin/collateral to mitigate the effects of default). Table 1 summarizes the key CFTC rulemaking in these areas.

The whole set of DFA regulations (not only the central clearing mandate) are likely to affect swaps trading. To identify the causal impact of the central clearing mandate, I need to examine a period when other regulations are not varying. The CFTC implemented the DFA regulations piecemeal during the 2012-2014 period, thus leaving only small windows where the impact of the clearing mandate can be studied in isolation (this is discussed further in the Identification Strategy section). I discuss how some of the regulations in Table 1 could impact trading. One of the provisions of the DFA that the CFTC implemented at the earliest was the data-reporting/record-keeping requirement. This required certain characteristics of swaps trades, such as the agreed upon rates and prices, to be reported in near real-time through SDRs. The OTC IR swaps market was previously relatively opaque (where quotes were usually obtained on a bilateral basis). The greater price transparency available to market participants after the implementation of the data-reporting regulation is likely to affect pricing and volatility (for example, see (Tarbert and Grimm 2021) which studies the impact of reporting requirement changes on swaps pricing).

The CFTC also encouraged standardization of swaps contracts by requiring parties to put up additional collateral for non-standard contracts, and for standardized contracts to be traded on (electronic) swaps execution facilities (SEF)/exchanges. SEFs are likely to increase competition (because requests for quotes are transmitted to multiple dealers simultaneously), increase pricing transparency (because it would give market participants access to price history, market depth and other market statistics) and increase liquidity (because it will allow more participants, both dealers and end-users to participate in the market) (Mateus and Faltoni 2015; McAlley 2024).

CFTC rulemaking also targeted the business conduct of swaps dealers and major swaps participants. This included requiring such entities to register with the CFTC, develop and maintain internal business conduct standards, set aside capital or require margining for trades they enter, segregate customer funds and have plans for unwinding trades in case of bankruptcy. These regulations are likely to reduce the risk/impact of a dealer default (counterparty risk) and to affect pricing and volatility through the reduced counterparty-risk channel.

Table 1 Major Rule-Making Areas of the Dodd-Frank Act.

The CFTC interpreted the DFA to contain six major rule-making areas (as well as a seventh, "other" area for miscalenous rules). Specific rules within each rule-making area are listed on the right (Commodity Futures Trading Commission, n.d.)

Rulemaking Area	Major Rules
Swaps Dealers and Major Swaps Participants	• Registration (Mar 2012) • Internal Business Conduct Standards (Jun 2012) • Capital and Margin for non-banks (2016; Sep-Nov 2020)[3] • Segregation (Apr 2012) and Bankruptcy (Feb 2012)
Data Requirements	• Data recordkeeping and reporting requirements (Oct 2010, Dec 2010, Mar 2012)[4] • Swap Data Repositories (SDR) Registration (Oct 2011) • Real Time Reporting (Mar 2012) • Large Swaps Trader Reporting (Sep 2011)
Clearing Requirements	• Establishment of Derivatives Clearing Organizations (DCO/CCP) (Jan 2012) • Clearing requirement for most common swaps (Mar-Sep 2013) • Margining requirements for uncleared swaps (Apr 2016)
Trading Requirements	• Establishment of Swaps Execution Facilities (SEF) (Jun 2013) • Made Available for Trade (MAT) designation/ requirement (Jun 2013)
Position Limits	• Position Limits and Aggregation of Positions (Jan 2012)
Enforcement	• Disruptive Trading Practices (Mar 2011) • Anti-Manipulation (Aug. 2011) • Whistleblowers (Aug 2011)
Other	• Reliance on Credit Ratings (July 2011) • Fair Credit Reporting Act (July 2011) • Investment Adviser Reporting (Nov 2011) • Volcker Rule (Jan 2014) • Cross-Border Applications (Jan 2013)

[3] The CFTC proposed capital and margin requirements starting in 2016. However these were implemented in phases, starting with the requirements applying to the largest SD and MSP. The final phase became effective in Nov. 2020.

[4] The main record-keeping and reporting requirements became effective in March 2012. The CFTC created interim rules in 2010 for record-keeping and reporting for "transition swaps", which are swaps that were created between the passage of the Dodd-Frank Act and the enactment of the final rules in 2012.

1.3.2 International Context

Considering the global nature of the financial system, regulators collaborated internationally to harmonize regulatory requirements. In Europe, the EU passed EMIR in 2012, which shared similar objectives as the DFA, while the Bank of England (BoE) issued regulations mandating clearing for most trades involving UK-based entities. In Asia, the Japanese Financial Services Authority (JFSA) required yen-denominated IR swaps and certain CD swaps to be cleared by the end of 2012; the Monetary Authority of Singapore (MAS) and the Securities and Futures Commission (SFC) of Hong Kong released consultation papers expressing their intentions to clear swaps denominated in certain Asian currencies. Table 2 summarizes clearing requirements internationally. Note that the table focuses only on the central clearing mandate in a global context, although other regulations (like those described in Table 1 for the US) were also enacted internationally as well.

The US and EU acted nearly simultaneously in enacting the central clearing requirement. As part of the EU, these regulations also affected trading in the UK (where London is a major financial center of swaps trading). In Japan and Australia, authorities enacted mandatory central clearing slightly before the US and EU, for contracts designated in their respective local currency. Other financial centers, such as Hong Kong, Singapore and Switzerland enacted central clearing requirements around the 2016-2017 period. Importantly, Canada implemented a central clearing requirement in May 2017, creating a period where IR swaps contracts denominated in Canadian dollars did not need to be cleared either in the US or in Canada.

Table 2 Summary of Central Clearing Requirements in Major Financial Centers.

Japan and Australia required mandatory clearing of JPY-denominated and AUD-denominated contracts traded in their jurisdictions starting at end of 2012. The US and EU required mandatory clearing (for contracts in various currencies) starting in the first half of 2013. Other countries enacted similar requirements between 2013 and 2017 .

Jurisdiction	**Relevant Laws and Regulations**
North America	• DFA (2010) and CFTC and SEC rulemaking requires mandatory clearing of IR swaps contracts denominated in USD LIBOR, GBP LIBOR, EURIBOR and JPY LIBOR by September 2013. • Additional currencies and classes of contracts are added to the clearing requirement in 2016 to harmonize regulations across jurisdictions. • Canada requires certain CAD-denominated swaps to be cleared starting in May 2017.
Europe	• EMIR passes in 2012 and requires clearing of certain IR swaps contracts. Regulations come into effect in March 2013. • Bank of England releases financial market regulatory guidance in April 2013, reiterating the applicability of EMIR to UK-based traders. • Additional currencies and classes of swaps are added to the EU clearing requirements in 2016. • Switzerland established a clearing mandate for Switzerland based swaps in 2017.
Asia	• Japan Financial Stability Authority (JFSA) requires JPY denominated IR swaps referencing JPY LIBOR to be cleared by end of 2012. • Hong Kong requires HKD denominated swaps to be cleared starting July 2017. • MAS requires SGD contracts to be cleared by December 2017.
Australia	• Australian Council of Financial Regulators (CFR) pass legislation requiring mandatory clearing of Australian dollar (AUD) denominated IR swaps by end of 2012.

1.4 Review of Literature

1.4.1 Interest Rate Swaps

Formal swap agreements were first seen in financial markets in 1981/1982. Bicksler and Chen (1986) find three uses of IR swaps in the market: (1) to manage mismatches in assets and liabilities (for example, depository institutions hold long-term fixed rate assets such as mortgages and short-term liabilities such as demand deposits; on the other hand, insurance companies often invest in short term assets that pay a variable rate, such as money market funds, and have long-term fixed-rate liabilities); to lower fixed-rate borrowing costs (borrowers with poor credit can often borrow at a lower cost in the floating rate market) and to manage their debt mix. The primary economic rationale for the existence of IR swaps is differences between firms' costs to borrow at fixed vs. variable rates arising due to market imperfections (differences in regulations or credit market imperfections can give firms comparative advantage in borrowing in one market over another).

Smith, et. al. (1988) present two models of pricing swaps. One model replicates the payoff of a swap through a portfolio of forward or futures contracts. The other model replicates the payoff through a portfolio of floating rate and fixed rate corporate bonds. They note that for a portfolio of bonds, there is an exchange of the principal at the end of the bond term, while for an IR swap the principal is usually not exchanged (that is, it is a "notional" principal). Thus, the impact of a default is greater for a corporate bond than for an IR swap. Futures contracts on the other hand are exchange-traded, cleared, and settled daily, so the risk of loss due to counterparty default is close to zero. For forwards, the contract value is realized only at the end of the contract period and has greater potential for counterparty default than for futures. An IR swap is somewhere in-between: it is periodically settled (on the payment dates).

Minton (1997) examines these valuation models. He finds that the fixed rate of the IR swap is discounted by ~4 bps compared to a replicating portfolio of Eurodollar futures (Eurostrips) and that movements in swap rates and Eurodollar futures rates are highly correlated. When evaluating the portfolio of bonds model, he finds that actual swap rates fall between the rate derived from a portfolio of corporate bonds and the rate derived from Eurodollar futures. Proxies for counterparty credit quality also have a statistically significant explanatory power, suggesting counterparty risk is a factor in observed swaps pricing.

1.4.2 Liquidity

Biais (1993) proposes a model for a dealer intermediated market and derives the optimal bid-ask spreads quoted by dealers with constant absolute risk aversion (CARA) (note that this is a model of general asset pricing in dealer-intermediated markets, and not specific to the IR swap market). A basic version of Biais' model motivates the liquidity model in section 2.2. Several papers empirically examine liquidity in the IR swap market: Sun, et. al. (1993) examine the effect of dealer credit rating on bid-ask spreads using data from Merrill Lynch and AIG Financial Products. They find that dealers with AAA credit ratings charge a spread of around ~10 bps while lower rated dealers only charge a spread of ~4 bps. Boudiaf, et. al (2024) examine the impact of monetary policy tightening on liquidity of EUR denominated swaps using a variety of liquidity measures. They find that their liquidity measures are impacted by monetary policy (specifically volatility in key policy rates reduces liquidity in the swaps market). Liu, et al (2006) decompose the "spread" between IR swaps and corresponding treasury bills into a credit spread and liquidity spread (arising from the lower liquidity of swaps over US government bonds). They find that the credit component of the spread is ~31 bps while the liquidity component is ~7 bps. Benos, et. al (2020) examine the impact of another Dodd-Frank mandate (trading on SEFs) on liquidity. They find a 12%-19% improvement in liquidity in the post-regulation market, driven by competition among dealers. Loon and Zhong (2014) examine liquidity in the credit default swap (CDS) market following the passage of the Dodd-Frank Act. They find that central clearing in the CDS market is associated with more liquidity.

1.4.3 Price Volatility

Compared to studies of pricing and liquidity, studies of price volatility in the IR swap market are rare. Azad, et. al (2012) decompose volatility in the US and UK market into high-frequency and low-frequency components using asymmetric spline GARCH (AS-GARCH). They then regress the low frequency component of the volatility against several macroeconomic variables (volatility of consumer price index, volatility of industrial production, volatility of short-term interest rates, volatility of foreign exchange rates, slope of the term-structure, unemployment rate and money supply). They find that volatility of short-term interest rates affects IR swap price volatility. In addition, for GBP based contracts, the money supply is negatively associated with IR swap volatility. For USD based contracts, the volatility of industrial production and the slope of the yield curve also affects IR swap volatility.

1.4.4 Systemic Risk and Contagion

Jackson and Pernoud (2021) outline two main avenues of contagion (that is financial distress at one institution spreading throughout the financial system): firstly, through defaults and fire-sales of assets that diminish the value of interconnected financial institutions (the network channel) and secondly, through feedback effects such as bank runs and credit freezes. For the first avenue, consider the case when a large financial institution fails. The values of other institutions that do business with the failing institution are also diminished and can cause a cascading series of failures. Each failure leads to additional bankruptcy costs and the final cost to the system at the end of the process can vastly exceed the size of initial shock. Such models are explored by Rochet & Tirole (1996) and Allen & Gale (2007; 2000).

Another way that financial institutions are interconnected are through the assets they trade. That is, even though two financial institutions might not directly do business with each other, they might own assets that are highly correlated. When a bank becomes insolvent, it often must sell assets at distressed prices. Such sales can also depress prices of related assets and drive institutions that hold those types of assets into insolvency. A prominent real-world example of this scenario is the 1998 crisis at Long Term Capital Management (LTCM) (Liu, Longstaff, and Mandell 2006). LTCM was a hedge fund that used a highly leveraged portfolio of IR swaps and foreign bonds (especially Russian bonds) to earn high market returns. When Russia defaulted on its debt in 1998 and devalued the Ruble, LTCM's portfolio took a large loss. In addition, market participants became more risk-averse and stopped lending to any institutions that employed a similar trading strategy to or held similar assets as LTCM. This created a system-wide credit crunch. The Federal Reserve eventually organized a bailout of the fund to prevent further damage to the financial system. Other prominent examples of this type of contagion are the Asian and Eurozone financial crises, where the potential default of one country led to distressed financial conditions in neighboring countries, as market participants became more risk averse. These types of models are explored by Kiyotaki & Moore (1997) , Cifuentes et al. (2005), Gai & Kapadia (2010), Capponi & Larsson (2015) and Greenwood et al. (2015).

Besides the network avenue, contagion can also occur through feedback loops and multiple equilibria. The classic Diamond and Dybvig (1983) model illustrates how multiple equilibria can lead to panic and bank runs. Banks lend out money long term and take in deposits short term. If enough depositors demand to withdraw their funds at once, the bank cannot repay all of them. In fact, if depositors believe a bank is insolvent (or they believe that other depositors believe that the bank is insolvent), they have an incentive to be the first in line to pull their funds out. Thus, a change in belief about the solvency of a bank can lead to a self-fulfilling insolvency, without any decrease in the value of the bank's actual portfolio of loans. Similarly, banks' beliefs about the creditworthiness of their counterparties can lead them to pull back their lending, leading to the very adverse credit condition and defaults that they were anticipating. This chain of defaults can cast doubts about the solvency of other banks, eventually leading to a systemwide freeze where banks stop lending to each other. These types of models are explored by Bebchuk & Goldstein (2011), Brunnermeier (2009) and Diamond and Rajan (2011).

1.4.5 Central Clearing

The policy and market implications of a central clearing mandate are discussed extensively by Pirrong (2011). Per Pirrong, CCPs should clear liquid, standardized products, as illiquid products can pose substantial risks to the CCP. CCP's can reduce the disruptive effect of defaults by drawing on additional sources of capital and facilitating orderly liquidation of positions. However, they can also increase systemic risk by requiring additional margin during periods of financial stress. In addition, by mutualizing the risk of default, they can induce market participants to take more risks (moral hazard and adverse selection issues). CCPs are also subject to economies of scale and scope (that is, the market will converge to one or few large CCPs that can economize over costs of warehousing and multiproduct netting). Since a CCP is likely to become a systemically important financial institution, regulators must monitor it closely and have prudent measures (such as a resolution plan if the CCP collapses).

Duffie and Zhu (2011) show that theoretically, concentrating clearing to one CCP can economize on collateral. Benos et al. (2019) explore the issue of economies of scale/scope among CCPs. Regulators in Europe and United States have required "local CCPs" to clear contracts that originate in their jurisdiction. They find that the same contracts trade at different prices when cleared through two different clearinghouses (LCH in the UK/Europe and CME Clearing in the US) and suggest that this difference arises due to increased collateral costs when clearing is fragmented.

Bernstein, et al. (2019) look at the impact of central clearing on equities pricing by examining the prices of the same stocks traded on New York Stock Exchange (NYSE) and Consolidated Stock Exchange (CSE). The NYSE established a clearinghouse in 1892 while the CSE did not. They find that the same stocks on the NYSE traded for 90-173 premium over the CSE price.

2 Theory

2.1 Model of Swaps Pricing

2.1.1 Pricing Without Credit Risks

An IR swap can be thought of as an exchange of a series of fixed payments by one party for a series of variable (floating) rate payments by the other party involved in the swap. For the fixed leg, the present value of the payments is given by (Darbyshire 2022):

$$PV_{fixed\ leg} = \sum_{i=0}^{T} \frac{CF}{(1+r_i)^{t_i}} \tag{1}$$

where: CF is the (fixed) cash flow, r_i is the risk-free rate for period i, t_i is the time at which CF will be received and T is the tenor (total length of the swap contract)

The present value of the floating leg is:

$$PV_{floating\ leg} = \sum_{i=0}^{T} \frac{CF_i}{(1+r_i)^{t_i}} \tag{2}$$

where: CF_i is the floating leg payment at period i, and all the other variables are as defined previously.

The present value of the contract for the party paying the fixed leg and receiving the floating leg is:

$$PV = PV_{floating\ leg} - PV_{fixed\ leg} \tag{3}$$

(The counterparty's value is given by a similar formula, but with the signs reversed on the right-hand side.)

Floating rate payments are unknown in advance but are usually forecasted by a relevant yield curve (Darbyshire 2022). For instance, if the floating leg payment is based on USD LIBOR, a USD

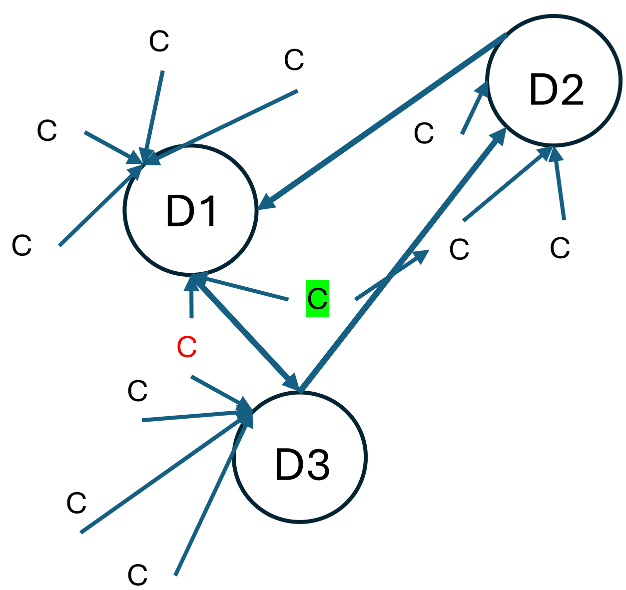
LIBOR curve, constructed by interpolating short-term deposit rates, medium-term Eurodollar futures, and long-term instruments like forward rate agreements and existing swaps, is used (Bloomberg L.P. 2024). At the outset of the contract, its value (PV) is zero. This is achieved by determining the present value of the floating leg using the forecasted payments (e.g. using the USD LIBOR yield curve) and then setting the fixed rate payment CF in (3) such that the present values of both legs equal. The payments are discounted using the same LIBOR yield curve.

2.1.2 Pricing with Counterparty Risk (Credit/Debit Valuation Adjustment)

The IR swap market is dominated by a handful of substantial swap dealers (SDs) and Major Swap Participants (MSPs) rather than many atomistic market participants (Bolandnazar 2020). These SDs and MSPs provide buy (bid) and sell (ask/offer) quotes for swaps, potentially finding other participants to balance their swap exposures. Figure 2 depicts a hypothetical network model of such a market.

Figure 2 A dealer-intermediated market witouth central clearing.

Three dealers (labeled D1 through D3) trade with many customers (labeled C). Dealers can engage in interdealer trades (shown with thicker arrows), and customers can trade with multiple dealers (such a customer is highlighted in red) or with other customers directly (such a customer is highlighted in green).



In the figure, three dealers (D) each engage with their set of clients (C). Note that dealers might engage in interdealer trading (indicated by thicker arrows between dealers) and bulk futures markets trading (not shown) for cash flow or risk management purposes. Customers can trade with multiple dealers (indicated by arrows going from C to multiple Ds) or occasionally engage in bilateral trades amongst themselves (indicated by arrows going from C to C). However, bilateral trades typically have low volume. Dealer-centric network structure lowers search costs compared to a direct customer-to-customer market (Bolandnazar 2020).

In practice, customers and dealers must account for the risk associated with counterparties defaulting. The "risk-free" present value pricing in equations (1)-(3) needs to be adjusted for this risk. If S_i represents the survival probability of the counterparty at period i, the expected present value of the fixed leg is:

$$PV = \sum_{i=0}^{T} \frac{CF_i \cdot S_i}{(1 + r_i)^{t_i}} \tag{4}$$

The fixed rate payment CF needs to account for the modified PV of the floating leg.

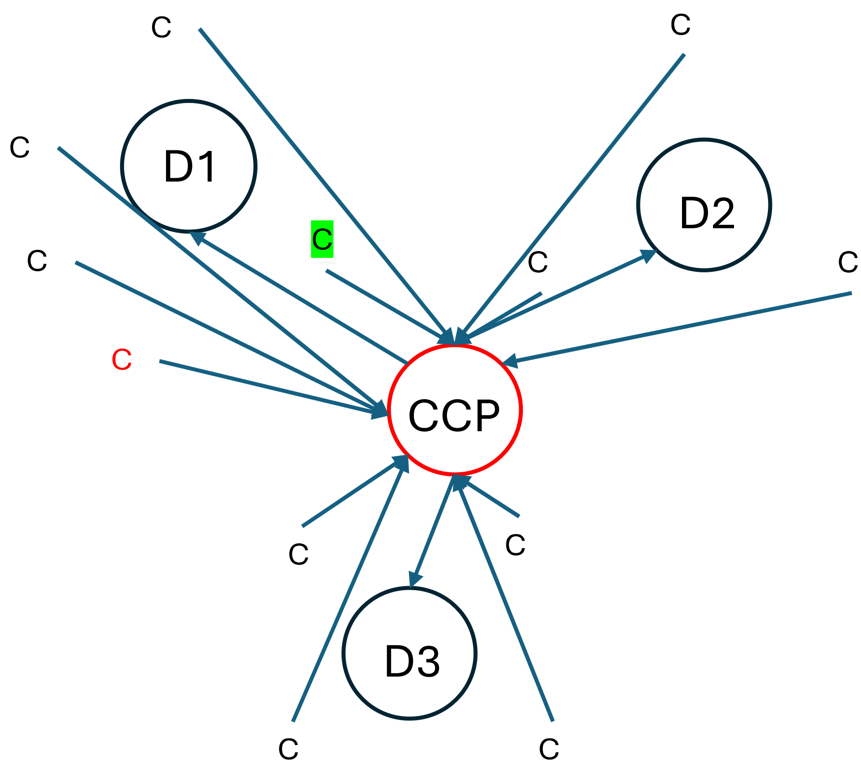
(Note that a swap's valuation with counterparty risk requires two adjustments. Only the credit valuation adjustment [CVA] is shown above. However, if one’s counterparty defaults, one no longer has to make their obligated payments to the other party either, which would increase the value of the contract. This adjustment is called the Debit Value Adjustment [DBA] and not shown above).

2.1.3 Pricing Under Central Clearing

The structure of a dealer-dominated market means that a dealer's failure (possibly due to inadequate risk management or correlated customer defaults) could affect other dealers and potentially the entire market. To counter this, regulators introduced central counterparties (clearinghouses). These clearinghouses void (novate) the initial swap contract and establish two new contracts, mirroring the original, with each counterparty. Now participants only need to be concerned about the clearinghouse's potential default, rather than their counterparties. Owing to their robust capitalization, regulation, and sound risk management, clearinghouses are perceived to decrease default and contagion risks. Figure 3 visualizes a hypothetical market structure with mandated central clearing. In this picture, the bilateral obligations between dealers (D) and customers (C) have been replaced by contracts between the dealer, customer, and the CCP.

Figure 3 Dealer-based Market with Central Clearing.

The obligations between customers and dealers, customers and customers, and between dealers are replaced by obligations between the CCP and the counterparty (dealer/customer).



If clearinghouses can reduce or eliminate counterparty risk, swaps values should be closer to the risk-free case rather than the case with counterparty risk. However, even if clearinghouses are successful at eliminating counterparty risk, additional cost of compliance (such as clearing fees and margin requirements) could keep swaps prices from reaching the risk-free valuation.

2.2 *Model of Liquidity (Bid-Ask Spreads)*

2.2.1 Liquidity with no Counterparty Risk

I adapt the model from Biais (1993) for centralized trading. I preserve the essential relationships between risk-aversion, volatility, dealers' inventory levels and the observed bid-ask spread but simplify the model to ease exposition. In the original model, traders (both liquidity traders and dealers) face trading costs. In addition, each market participant faces a fixed cost if they choose to become a dealer (these costs could be due to efforts dealers need to expend to monitor the market, maintain a presence on the trading floor, engage in risk-control and back-office activities). At the start of the sequential game outlined by Biais, M of the N market participants choose to become dealers based on their costs and expected profits from market-making activities. However, this aspect of Biais' model is not consequential to the fundamental relationship between the bid and ask prices quoted by dealers and the volatility of underlying prices. To simplify exposition, I modify the sequential game to involve exactly 2 dealers and one liquidity trader (that is, players no longer choose their type but are predetermined to either be dealers or a liquidity trader). In addition, I drop the fixed trading costs for dealers, as this only adds a constant level to dealers' reservation quotes.

The model is a sequential game where all market participants (players) can observe the quotes (bids and asks) and (market) orders of market participants. In the first stage, two competing dealers (liquidity providers) receive a random inventory position in the risky asset between [-R, R]. In stage 2, dealers set their bid and ask prices. In the next stage there is a liquidity shock with probability λ. If there is a liquidity shock, a liquidity trader (liquidity demander) receives an inventory of quantity L with probability $\frac{1}{2}$ (or a short position of size $-L$ with the same probability). They then decide the size of the optimal market order, which is executed at the best bid or ask price posted by the dealers. If there is no liquidity shock, no trade takes place. In the final stage, the price of the risky asset is realized, and players receive their utility.

I assume that the two dealers are identical except for their inventory positions. All market participants have constant absolute risk aversion (CARA) utility $U(w) = -\exp(-\alpha \cdot w)$, where w is the trader's wealth and α is a risk-aversion parameter. If there is a liquidity shock, the liquidity trader observes the market prices and selects the quantity (size) of the market order. The final price of the risky asset is $p = 1 + z$, where $z \sim \mathcal{N}(0, \sigma^2)$.

I analyze the case where the liquidity trader receives a liquidity shock $+L$ (the case where they receive a $-L$ shock will be analogous). At the end of the game, once the asset price is realized, the trader receives wealth:

$$w = p(L - Q) + Q \cdot p_b^* \tag{5}$$

where $L - Q$ is the net position of the trader in the risky asset at the end of the game and $(Q \cdot p_b^*)$ is the cash from selling Q units of the risky asset at the best (highest) bid price p_b^*. The liquidity trader will maximize expected utility $\mathbb{E}[U(w)] = \mathbb{E}[-\exp \alpha\big((L - Q)p + Q \cdot p_b^*\big)]$. The optimal quantity is:

$$Q^* = \frac{1}{a \cdot \sigma^2}(p_b^* - 1) + L \tag{6}$$

where I have used the fact that $\mathbb{E}[\exp(z)], z \sim \mathcal{N}(\mu, \sigma^2) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$ and $\mu = 0$.

If dealer 1 has the best price, they receive the order flow and have wealth:

$$w = (I_1 + Q^*)p - p_1^b Q \tag{7}$$

where: p_1^b is the bid price set by dealer 1, Q^* is the size of the market order and I_1 is their random inventory position (a number between $[-R, R]$). If dealer 1 does not have the best price, they receive:

$$w = I_1 p$$

Dealer 1 is indifferent between trading and not trading when the expected utility from both actions is the same. This happens when:

$$p_{r,1}^b = \frac{\alpha\sigma^2}{2}(Q^* + 2I_1) \tag{8}$$

$$p_{r,1}^b = \alpha\sigma^2(L + 2I_1) - 1 \tag{9}$$

where the price is subscripted with r to emphasize it is the reservation price. A similar analysis holds for the ask price (when dealers are competing over market sell orders in response to a negative liquidity shock):

$$p_{r,1}^a = \alpha\sigma^2(L - 2I_1) - 1 \tag{10}$$

and in general, for dealer i, the reservation prices are:

$$p_{r,i}^b = \alpha\sigma^2(L + 2I_i) - 1 \tag{11}$$

$$p_{r,i}^a = \alpha\sigma^2(L - 2I_i) - 1$$

I analyze the case of the optimal bid quote for dealer 1, assuming the competing dealer does not observe the other's inventory level, but both assume that the other's inventory is drawn uniformly from $[-R, R]$. A dealer can increase their probability of winning the order flow by improving their bid price but must balance this against the fact that they pay more for each unit acquired. The dealer would not like to increase the quote beyond their reservation price. The optimal bid quote for dealer 1 is:

$$p_1^b = p_{r,1}^b + \alpha \cdot \sigma^2 \frac{(R - I_1)}{2} \tag{12}$$

Similarly, the optimal ask quote is:

$$p_1^a = p_{r,1}^a + \alpha \cdot \sigma^2 \frac{(R + I_1)}{2} \tag{13}$$

and in general, the optimal bid and ask quotes for dealer i are:

$$p_i^a = p_{r,i}^a + \alpha \cdot \sigma^2 \frac{(R + I_i)}{2} \tag{14}$$

$$p_i^b = p_{r,i}^b + \alpha \cdot \sigma^2 \frac{(R - I_i)}{2}$$

The observed bid ask spread is:

$$S = \max[p_i^b] - \min[p_i^a] \tag{15}$$

Note: Under competition with N dealers, the second term on the RHS of (14) becomes $\left[\alpha \cdot \sigma^2 \frac{(R \pm I_i)}{N}\right]$ and approaches 0 as $N \to \infty$. The bid-ask quotes collapse to the reservation quotes.

2.2.2 Sensitivity to Specification of Utility Function

In the above, I derived expressions for the bid-ask spread using a Constant Absolute Risk Aversion (CARA) utility function for all market participants. In this section, I show that when the utility function is switched to Constant Relative Risk Aversion (CRRA) utility, the optimal bid-ask spreads are qualitatively similar.

Consider the same game-theoretic setup described previously. However, now players have utility functions:

$$u(w) = \begin{cases} \frac{w^{1-\gamma}}{1-\gamma}, \gamma \neq 1 \\ \ln w, \gamma = 1 \end{cases} \tag{16}$$

The final wealth of the liquidity trader is given by:

$$w(Q) = Q \cdot p_b + (L - Q) \cdot (1 + z) = Q(p_b - 1) + (L + Q)z \tag{17}$$
$$z \sim \mathcal{N}(0, \sigma^2)$$

Let $\mu_w = \mathbb{E}[w(Q)]$ be the expected final wealth. If we expect the size of the liquidity shock and order size to be relatively small, we can approximate the utility function around μ_w by a second-order Taylor series approximation:

$$u(w) \approx u(\mu_w) + (w - \mu_w)u'(\mu_w) + \frac{1}{2}(w - \mu_w)^2 u''(\mu_w) \tag{18}$$

then

$$\mathbb{E}[u(w)] \approx u(\mu_w) + \frac{1}{2}u''(\mu_w)\mathbb{E}[(w - \mu_w)^2] \tag{19}$$

where I have used $\mathbb{E}[w - \mu_w] = 0$

Since $\mathbb{E}[(w - \mu_w)^2] = (L - Q)^2\sigma^2$. We have:

$$\mathbb{E}[u(w)|w = \mu_w] \approx \frac{\mu_w(Q)^{1-\gamma} - 1}{1 - \gamma} - \frac{\gamma}{2}\mu_w(Q)^{-\gamma-1}(L - Q)^2\sigma^2 \tag{20}$$

where I write $\mu_w = \mu_w(Q)$ to emphasize that the trader's expected wealth is a function of their order quantity. To find the optimal order quantity, we can set the derivative of the expected utility to zero:

$$d\frac{\mathbb{E}[u(w(Q))]}{dQ} = 0 \tag{21}$$

We can solve this to obtain Q*:

$$Q^* = L + \frac{(p_b - 1)\mu_w(Q)^{2\gamma+1}}{\gamma\sigma^2} \tag{22}$$

This is similar the optimal order quantity we derived under CARA (see equation (6)) with an additional $\mu_w^{2\gamma+1}$ term in the numerator, showing the order quantity is sensitive to the expected level of wealth of the trader.

We can similarly derive the reservation prices for the dealers. For simplicity, I work through the reservation bid price for dealer 1.

Dealer 1 receives a random inventory position I_1, $I_1 \sim Unif(-R, R)$. If they post the best bid price and receive the trade and their wealth at the end of the period is:

$$w = (I_1 + Q)(1 + z) - p_b^1 Q \tag{23}$$

If they do not have the best price, they do not trade, and their wealth will be:

$$w^{NT} = I_1(1 + z) \tag{24}$$

At some price $p_{b,r}^1$ the dealer will be indifferent between trading and not trading. This will occur when $\mathbb{E}\left[u\left(w(p_{b.r}^1)\right)\right] = \mathbb{E}[u(w^{NT})]$.

Let μ_w be the expected wealth $\mathbb{E}\left[(1 + Q)(1 + z) - p_{b,r}^1 Q\right] = I_1 + Q\left(1 - p_{b,r}^1\right)$

The second order Taylor-series expansion about μ_w is:

$$u(w)|w = \mu_w \approx u(\mu_w) + (w - \mu_w)u'(\mu_w) + \frac{1}{2}(\mu_w - w)^2 u''(\mu_w) \tag{25}$$

Taking expectations:

$$\mathbb{E}[u(w)|w = \mu_w] \approx \mathbb{E}[u(\mu_w)] + \frac{1}{2}(w - \mu_w)^2 u''(\mu_w) \tag{26}$$

Using the fact that $\mathbb{E}[w - \mu_w]^2 = (I_1 + Q)^2\sigma^2$ we obtain:

$$\mathbb{E}[u] \approx \frac{I_1 + Q\left(1 - P_{b,r}^1\right)^{1-\gamma} - 1}{1-\gamma} + \frac{1}{2}(I_1 + Q)^2\sigma^2\left(-\gamma(I_1 + Q\left(1 - p_{b,r}^1\right)^{-\gamma-1}\right. \tag{27}$$

We can similarly take the Taylor Series approximation of the utility function at the no-trade level of wealth:

$$\mathbb{E}[u(w_{NT})] \approx \frac{\left(w_{NT}^{1-\gamma} - 1\right)}{1-\gamma} - \frac{\gamma}{2} I_1^2 \sigma^2 w_{NT}^{-\gamma-1} \tag{28}$$

We can substitute $w_{NT} = I_1(1 + z)$ and $w = (I_1 + Q)(1 + z) - p_b^1 Q$ into the expressions above and set $\mathbb{E}[u(w)|w = \mu_w] = \mathbb{E}[u(w_{NT})]$. We obtain the reservation price:

$$p_{b,r}^1 \approx \frac{\gamma\sigma^2}{2}\mu_w(p_{r,b}^1) \tag{29}$$

We cannot solve for $p_{b,r}^1$ explicitly, but note it has a similar form to the CARA reservation price.

In competitive markets, the dealer's markup will tend to zero and prices will tend to the reservation prices.

Thus, we find that using CRRA utility function (or with an utility function which is at least as concave as a CRRA utility function), the optimal order quantity and bid-ask spreads are qualitatively similar to the CARA utility case, except quantities and spreads are sensitive to the level of expected wealth of the market participants.

2.2.3 Liquidity with Counterparty Risk

Under the scenario where there is counterparty risk, if the counterparty defaults the value of the asset is impaired (the holder of the non-defaulting leg no longer receives expected cash flows). However, for the defaulter, the value of the asset is enhanced (as they no longer need to make payments). I model this as an additional shock to the realized value of the asset: $p = (1 + z + y), y \sim \mathcal{N}(0, \delta^2)$. The analysis remains essentially the same, but the optimal market order size, reservation quotes, and optimal quotes now become:

$$Q^* = \frac{1}{a \cdot (\sigma + \delta)^2}(p_b - 1) + L \tag{30}$$

$$p_{r,i}^b = \alpha(\sigma + \delta)^2(L + 2I_i) - 1 \tag{31}$$

$$p_{r,i}^a = \alpha(\sigma + \delta)^2(L - 2I_i) - 1 \tag{32}$$

$$p_i^a = p_{r,i}^a + \alpha \cdot (\sigma + \delta)^2 \frac{(R + I_i)}{2} \tag{33}$$

$$p_i^b = p_{r,i}^b + \alpha \cdot (\sigma + \delta)^2 \frac{(R - I_i)}{2} \tag{34}$$

The optimal order size is reduced from $Q^* \propto \frac{1}{\sigma^2}$ to $Q^* \propto \frac{1}{(\sigma+\delta)^2}$, the reservation quotes are increased from $p_r \propto \sigma^2$ to $p_r \propto (\sigma + \delta)^2$ and the mark-up from the reservation prices are similarly affected. With counterparty risk, the market order size is smaller, and the bid-ask spread is larger than under no counterparty risk.

Since Central Clearing is supposed to mitigate counterparty risk (and lower volatility, see next section), I expect observed bid-ask spreads under central clearing to be somewhere between the no counterparty risk and counterparty risk case.

2.3 Model of Price Volatility

2.3.1 Volatility without Counterparty Risk

I develop an original model of price volatility. There are two sets of agents: market-makers who post bid and ask prices, and liquidity traders who post market orders. Assume that order-flow (net market-buy or market-sell orders) is i.i.d Normal with variance σ_ϵ^2:

$$OF_t \sim \mathcal{N}(0, \sigma_\epsilon^2) \tag{35}$$

Market-makers adjust their next period price (P_{t+1}) based on the current period's observed price (P_t) and order-flow (OF_t):

$$P_{t+1} = P_t + \alpha \cdot OF_t \tag{36}$$

where: α is a parameter for the market-makers' sensitivity to order-flow.

The expression for the volatility in this case is:

$$vol = \sqrt{Var(P_{t+1} - P_t)} = \alpha \cdot \sigma_\epsilon \tag{37}$$

2.3.2 Price Volatility in Markets with Counterparty Risk

I modify the above model to include additional order-flow dynamics related to counterparty risk. Assume that when the current period's order-flow is negative, there is additional sell-off of the risky asset in the next period due to (perceived) additional counterparty risk, and when the current period's order flow is positive, there is additional buying of the risky asset in the next period due to (perceived) reduction in counterparty risk. The order-flow dynamics are now given by:

$$OF_{t+1} = \rho OF_t + \epsilon_{t+1} \tag{38}$$

Where: $\epsilon_{t+1} \sim \mathcal{N}(0, \sigma_\epsilon^2)$

The expression for the price change becomes:

$$\Delta P_t = (P_{t+1} - P_t) = \alpha OF_{t+1} \tag{39}$$

$$vol = \sqrt{Var(\Delta P_t)} = \alpha^2 Var(OF_{t+1}) \tag{40}$$

$$Var(OF_t) = \frac{Var(\epsilon_{t+1})}{1 - \rho^2} = \frac{\sigma_e^2}{1 - \rho^2} \tag{41}$$

$$vol = \alpha \sigma_\epsilon \sqrt{\frac{1}{1 - \rho^2}} \tag{42}$$

where I have used the fact $Var(OF_{t+1}) = Var(OF_t)$ on the 3rd line and that $Var(\epsilon_{t+1}) = \sigma_\epsilon^2$.

2.4 *Summary of Theoretical Findings*

I summarize here my expectations of what I expect to find in the empirical section from my theory:

1. Pricing: I expect the prices/premia for swaps to increase in the post-clearing period due to lower counterparty risk via the demand channel. That is, at any given price, the demand for

IR swaps in the post-clearing period will be greater since counterparty default risk is now lower.

2. Liquidity: I expect a narrowing of bid-ask spreads in the post-clearing period due to a decline in the unobservable δ term in equations (33)-(34). This term represents "extra volatility" to prices due to counterparty defaults and should be lower (or zero) if central clearing reduces (eliminates) counterparty risk.
3. Volatility: I expect a lower volatility in the post-clearing period through a reduction in the $\sqrt{\frac{1}{1-\rho^2}}$ term in (42). I expect $\rho \to 0$ (the no counterparty risk case as in (36)) as counterparty risk decreases, which would reduce the observed volatility.

3 Data

3.1 Data Sources and Collection

The primary dataset used in this analysis consists of trade-level information on IR swaps obtained from Bloomberg's Swaps Data Repository (SDR) screen. Bloomberg compiles and disseminates this data from the Depository Trust and Clearing Corporation's Swaps Data Repository (DTCC SDR), one of the largest repositories designated by the Commodity Futures Trading Commission (CFTC) to collect and maintain records of swap transactions.

The data collection process is governed by the reporting requirements introduced under the DFA. Specifically, the CFTC mandated that swap counterparties report detailed transaction-level data to registered SDRs, such as DTCC, shortly after trade execution (*Swap Data Recordkeeping and Reporting Requirements* 2012). These reporting obligations, which became effective starting in late 2012, aimed at increasing market transparency and improving regulatory oversight by providing near real-time access to key trade characteristics.

The DTCC SDR dataset from Bloomberg captures detailed trade information, including trade date and time, swap currency, notional value, fixed and floating rates, contract tenor, payment frequency, capped notional indicators, and clearing status. The dataset spans the windows around the phased implementation of the CFTC's central clearing mandate, providing comprehensive coverage for analyzing the causal impact of the regulation on pricing, liquidity, and volatility of IR swaps.

In order to price swaps properly, I need relevant yield curves. I also obtain the data for this from the Bloomberg Terminal (YC function). I detail the exact curve building methodology (including the specific curves used) in a later section.

3.2 Raw Data Description

The raw dataset comprises individual trade-level observations of IR swaps from the DTCC SDR, accessed via Bloomberg's terminal. Each trade observation includes detailed ~~characteristics necessary for comprehensive analysis~~:

- **Trade Date and Time**: Timestamp indicating when each swap trade was executed.
- **Swap Currency**: Denomination currency of the swap (primarily USD and CAD in this study).
- **Notional Value**: The principal amount used to calculate payments exchanged by counterparties.
- **Fixed and Floating Rates**: The fixed interest rate agreed upon in the swap contract, and the reference rate for floating payments (e.g., LIBOR for USD swaps and CDOR for CAD swaps).
- **Contract Tenor**: Duration of the swap contract from effective date to maturity.
- **Payment Frequency**: Interval at which payments are exchanged (semiannual for fixed leg, quarterly for floating leg).
- **Capped Notional Indicators**: Indicators specifying if swaps have capped notional amounts.
- **Clearing Status**: Indicator specifying whether a swap was centrally cleared or uncleared.

The dataset covers trades conducted around the three implementation phases of the CFTC's central clearing mandate: Phase 1 (March 11, 2013), Phase 2 (June 10, 2013), and Phase 3 (September 9, 2013). These periods include ten trading days before and after each implementation phase, ensuring comparability across pre- and post-regulatory environments.

Table 3 shows the counts and total notional values by floating leg reference. LIBOR was the most common floating leg reference for USD denominated contracts, used for more than 95% of all USD contracts (whether by count or total notional value). The CDOR was the most common reference rate for CAD-denominated contracts (again for more than 95% of CAD-denominated contracts). Other reference rates for USD contracts included the Fed Funds Rate, Prime Rate, OIS Rate, and several municipal rate indices. Interestingly, there are several contracts that use a foreign reference rate (COOBIVR is the Colombian overnight rate, IBR is the Colombian equivalent of LIBOR, CLP-TNA is a Chilean rate, CLICP is a Chilean price index). These are likely contracts that are hedging or speculating foreign interest rate or inflation risk but want to be paid out in USD. For CAD-denominated contracts, besides the CDOR, the other reference rate was the CORRA (which is an overnight rate). Clearing trended from 61% prior to implementation of phase 1, to 78% after phase 1, and 89% after phase 2. It stayed at 89% following the implementation of phase 3. For CAD-denominated swaps, clearing hovered between 48%-59%.

Table 3 Number of contracts and notional values by clearing status and reference rate for USD and CAD IR swaps (unfiltered data set).

Data are presented for each phase separately. Within each phase, the pre-implementtion and post-implementation period are reported separately.

Pre phase 1 implementation (Feb 25 – Mar 8)						
CUR	**Floating Leg**	**Cleared (Count)**	**Cleared (Notional Value)**	**Uncleared (Count)**	**Uncleared (Notional Value)**	**Percent Cleared**
USD	LIBOR	3,518	203,345.90	3,071	131,242.01	61%
	USD-Federal Funds-H.15	0	0.0	16	2,183.00	0%
	USD-PRIME-H.15	0	0.0	2	6.00	0%
	USD-PRIME-H15	0	0.0	2	4.00	0%
	USD SPRDL MANUAL	0	0.0	1	100.00	0%
	USD-AAA_MUNI-	0	0.0	4	31.00	0%
	USD-OIS-3	0	0.0	1	6.00	0%
	IBR	0	0.0	2	200.00	0%
	CLICP	0	0.0	1	100.00	0%
	TIS	0	0.0	1	1.00	0%
	USD-USPSA-BLOOMBERG	0	0.0	1	4.00	0%
CAD	CAD-BA-CDOR	225	18,811.40	308	20,363.10	48%
	CAD-REPO-CORRA	0	0.00	3	410.00	0%

Post phase 1 implementation (Mar 11 – Mar 22)						
CUR	**Floating Leg**	**Cleared (Count)**	**Cleared (Notional Value)**	**Uncleared (Count)**	**Uncleared (Notional Value)**	**Percent Cleared**
USD	LIBOR	4,342	262,257.70	2,125	76,649.65	77%
	USD-Federal Funds-H.15	0	0.00	24	3,353.00	0%
	IBR	0	0.00	6	1,050.00	0%
	USD-SIFMA Municipal Swap Index	0	0.00	6	60.00	0%
	USD-PRIME-H.15	0	0.00	2	6.00	0%
	USD-PRIME-H15	0	0.00	2	3.00	0%
	USD-Prime-H.15	0	0.00	1	2.00	0%
	USD-USPSA-BLOOMBERG	0	0.00	2	20.00	0%
	CLICP	0	0.00	3	450.00	0%
	USD-AAA_MUNI-	0	0.00	2	25.00	0%
	USD-BMA Municipal Swap Index	0	0.00	2	6.52	0%
CAD	CAD-BA-CDOR	126	9,578.00	140	11,137.31	46%
	CAD-REPO-CORRA	0	0.00	3	780.00	0%
	CDOR	0	0.00	5	105.60	0%

Pre phase 2 implementation (May 27 – Jun 7)						
CUR	**Floating Leg**	**Cleared (Count)**	**Cleared (Notional Value)**	**Uncleared (Count)**	**Uncleared (Notional Value)**	**Percent Cleared**
USD	LIBOR	6,870	426,753.26	2,954	118,388.28	78%
	USD-Federal Funds-H.15	0	0.00	29	4,463.00	0%
	COOVIBR	0	0.00	9	1,800.00	0%
	CLP-TNA	0	0.00	6	1,200.00	0%
	USD FORM 3750	0	0.00	1	100.00	0%
	USD-AAA_MUNI-	0	0.00	2	13.00	0%
	USD-BMA Municipal Swap Index	0	0.00	1	7.00	0%
	USD-PRIME-H.15	0	0.00	12	62.90	0%
CAD	CAD-BA-CDOR	180	14,726.00	169	14,290.70	51%

Post phase 2 implementation (Jun 10 – Jun 21)						
CUR	**Floating Leg**	**Cleared (Count)**	**Cleared (Notional Value)**	**Uncleared (Count)**	**Uncleared (Notional Value)**	**Percent Cleared**
USD	LIBOR	7,975.00	461,124.51	1,449.00	53,548.33	90%
	USD-Federal Funds-H.15	0.00	0.00	33.00	5,068.00	0%
	USD-PRIME-H.15	0.00	0.00	5.00	26.00	0%
	USD-PRIME-H15	0.00	0.00	1.00	9.00	0%
	USD-Prime-H.15	0.00	0.00	1.00	2.00	0%
	COOVIBR	0.00	0.00	21.00	3,750.00	0%
	CLP-TNA	0.00	0.00	7.00	700.00	0%
	USD BMA MANUAL	0.00	0.00	1.00	45.00	0%
	USD-AAA_MUNI-	0.00	0.00	3.00	20.00	0%
	USD-BMA Municipal Swap Index	0.00	0.00	2.00	10.00	0%
CAD	CAD-BA-CDOR	176.00	11,322.08	174.00	7,969.50	59%

Pre phase 3 implementation (Aug 26 – Sep 6)						
CUR	**Floating Leg**	**Cleared (Count)**	**Cleared (Notional value)**	**Uncleared (Count)**	**Uncleared (Notional Value)**	**Percent Cleared**
USD	LIBOR	6,112.00	396,744.28	1,398.00	47,355.82	89%
	USD-Federal Funds-H.15	0.00	0.00	36.00	4,539.00	0%
	USD-PRIME-WEIGHTED-AVERAGE	0.00	0.00	2.00	200.00	0%
	USD-PRIME-H.15	0.00	0.00	5.00	7.00	0%
	USD-PRIME-H15	0.00	0.00	8.00	35.56	0%
	USD-AAA_MUNI-	0.00	0.00	1.00	10.00	0%
	USD-SIFMA Municipal Swap Index	0.00	0.00	1.00	5.00	0%
CAD	CAD-BA-CDOR	128.00	9,697.20	134.00	7,487.11	56%
	CAD-REPO-CORRA	0.00	0.00	1.00	35.00	0%

Post phase 3 implementation (Sep 9 – Sep 20)						
CUR	**Floating Leg**	**Cleared (Count)**	**Cleared (Notional Value)**	**Uncleared (Count)**	**Uncleared (Notional Value)**	**Percent Cleared**
USD	LIBOR	7,481	485,507.61	1,461	58,912.20	89%
	USD-Federal Funds-H.15	0	0.00	19	3,606.00	0%
	TREASURY_DTCC_GCF_REPO_INDEX	0	0.00	4	850.00	0%
	USD FORM 3750	0	0.00	1	30.00	0%
	USD-AAA_MUNI-	0	0.00	9	56.00	0%
	USD-BMA Municipal Swap Index	0	0.00	3	13.00	0%
	USD-BMA-BMA	0	0.00	1	22.00	0%
	USD-BMA-REFB	0	0.00	2	12.75	0%
	USD-PRIME-H.15	0	0.00	7	17.00	0%
	USD-PRIME-H15	0	0.00	7	64.00	0%
	USD-Prime-H.15	0	0.00	1	1.00	0%
	USD-SIFMA Municipal Swap Index	0	0.00	5	52.75	0%
CAD	CAD-BA-CDOR	210	14,099.00	354	15,561.41	48%
	CDOR	0	0.00	1	5.00	0%
	CDOR.CAD	0	0.00	4	106.00	0%

3.3 Data Filtering and Cleaning

To ensure accuracy and comparability, several filtering and cleaning steps were applied to the raw dataset:

- **Premium Threshold**: Trades with a premium outside ±50 basis points (bps) relative to Bloomberg's fair valuation of a similar swap contract were excluded. These outliers likely involve unique or non-standard features not captured in the dataset.
- **Currency and Reference Rate Filtering**: Only contracts denominated in USD and CAD with USD LIBOR (for USD) or CDOR (for CAD) as the floating reference rates were retained, as these are most relevant for the study.
- **Contract Characteristics**: Single-payment (zero-coupon) swaps and contracts with non-standard payment frequencies were excluded due to significant pricing differences compared to standard contracts.
- **Voluntary Clearing Status**: Contracts that were voluntarily cleared before the central clearing mandate or remained uncleared due to specific exemptions after the mandate were excluded to isolate the causal impact of mandatory clearing.

In addition, the following data cleaning steps were taken: (1) if the original rates were expressed in basis points (detected by the rate being greater than 10), they were converted to percentages by dividing by 100 and (2) usually the `Rate 1` is the fixed leg and `Rate 2` is the variable leg in the dataset; in some records these legs are "flipped" and corrections were made to account for this.

~~These steps resulted in a refined dataset suitable for analyzing the central clearing mandate's impact on pricing, liquidity, and volatility of standardized IR swaps.~~

3.4 Yield Curve Construction

To calculate the theoretical counterparty-riskless price of IR swaps, I forecast future floating rate payments and discount the payments using the appropriate yield curve. I use a single curve method, the prevalent pricing method during the study period (subsequently, the market switched to a dual-curve method of pricing swaps, where one curve was used to calculate future floating-rate payments, and another curve to discount those payments to their present value). For USD swaps, I obtain the USD semiannual fixed-floating rate curve (curve S23) for each trading day from Bloomberg. I similarly obtain the Canadian yield curve (curve S11) from the Bloomberg Terminal for pricing Canadian swaps. I use curve S45 for EUR denominated contracts, curve S10 for GBP denominated contracts and S12 for CHF denominated contracts.

I use the QuantLib-python library to construct the forward curve (Ametrano and Ballabio 2003a). For the USD swaps curve, the short-end (3M or less) of the curve is anchored by LIBOR rates; the medium-end (6M – 18M) of the curve is anchored by Eurodollar futures; and the long-end (24M onward) of the curve is anchored by US swap rates.

Table 4 shows sample data for CAD and USD yield curves on September 11, 2013. Note that futures rates need to have a convexity adjustment applied to them since futures payoffs differ from payoffs for other instruments (Bloomberg L.P. 2024). The values reported in the table have this convexity adjustment applied. Values between the “pillars” (observed data points) of the yield curve need to be interpolated. I use piecewise linear interpolation. I verify the curve by pricing contracts using my constructed curve and comparing against calculations by Bloomberg SWPM function. I can match the output of Bloomberg’s SWPM up to 4 decimal places.

Table 4 Sample data for construction of USD and CAD swaps curves.

The "short end" of the curve is based on observed deposit rates (such as the overnight Canadian Call Loan Rate and the 1-month and 2-month Canadian Offer Rate). The "medium leg" of the curve is constructed using data from futures contracts (note that a convexity adjustment needs to be applied to quoted futures prices due to differences in settlement between swaps and futures) and the "long end" of the curve is constructed using data from observed IR swaps. (Bloomberg L.P. 2013)

Tenor	Bloomberg CUSIP	Yield	Data Source[5]
3M	EDU13 Comdty	0.2575	BGN
6M	EDZ13 Comdty	0.294	BGN
9M	EDH14 Comdty	0.3574	BGN
12M	EDM14 Comdty	0.4402	BGN
15M	EDU14 Comdty	0.5675	BGN
18M	EDZ14 Comdty	0.7341	BGN
2Y	USSWAP2 BGN Curncy	0.5957	BGN
3Y	USSWAP3 BGN Curncy	1.0014	BGN
4Y	USSWAP4 BGN Curncy	1.45	BGN
5Y	USSWAP5 BGN Curncy	1.865	BGN
6Y	USSW6 BGN Curncy	2.2145	BGN
7Y	USSWAP7 BGN Curncy	2.501	BGN
8Y	USSW8 BGN Curncy	2.7305	BGN
9Y	USSW9 BGN Curncy	2.919	BGN
10Y	USSWAP10 BGN Curncy	3.0765	BGN
11Y	USSWAP11 BGN Curncy	3.2103	BGN
12Y	USSWAP12 BGN Curncy	3.322	BGN
15Y	USSWAP15 BGN Curncy	3.551	BGN
20Y	USSWAP20 BGN Curncy	3.7315	BGN
25Y	USSWAP25 BGN Curncy	3.815	BGN
30Y	USSWAP30 BGN Curncy	3.8565	BGN

Note: USD Swaps Curve. Last Updated Date: 9/11/13

Tenor	CUSIP	Yield	Data Source[5]
1D	CCLR Index	1.00	CMPN
1M	CDOR01 Index	1.22	CMPN
2M	CDOR02 Index	1.2475	CMPN
3M	BAU13 Comdty	1.275	BGN
6M	BAZ13 Comdty	1.2997	BGN
9M	BAH14 Comdty	1.3491	BGN

[5] The "data source" column is the Bloomberg terminal reported methodology/source for the values in the yield column. BGN refers to Bloomberg Generic and CMPN refers to Composite of Real-Time Contributed Prices. The BGN source is derived from contributions from market participants, executable quotes, and Bloomberg's own pricing models. CMPN is based on actual contributions from dealers, such as bid/ask quotes and indicative prices. I do not cite BGN or CMPN in references.

12M	BAM14 Comdty	1.4584	BGN
15M	BAU14 Comdty	1.6275	BGN
18M	BAZ14 Comdty	1.8164	BGN
2Y	CDSW2 BGN Curncy	1.6195	BGN
3Y	CDSW3 BGN Curncy	1.9372	BGN
4Y	CDSW4 BGN Curncy	2.235	BGN
5Y	CDSW5 BGN Curncy	2.4855	BGN
6Y	CDSW6 BGN Curncy	2.6885	BGN
7Y	CDSW7 BGN Curncy	2.8595	BGN
8Y	CDSW8 BGN Curncy	3.003	BGN
9Y	CDSW9 BGN Curncy	3.1335	BGN
10Y	CDSW10 BGN Curncy	3.254	BGN
12Y	CDSW12 BGN Curncy	3.457	BGN
15Y	CDSW15 BGN Curncy	3.6713	BGN
20Y	CDSW20 BGN Curncy	3.7915	BGN
25Y	CDSW25 BGN Curncy	3.7555	BGN
30Y	CDSW30 BGN Curncy	3.693	BGN

Note: CAD Swaps Curve. Last Updated Date: 9/11/13

3.5 Descriptive Statistics

Table 5 shows summary statistics of the control variables used in the regressions (see sections 4 and 5). Wednesday was the most active trading day and Monday and Friday were the least active trading days. The dataset includes two trading holidays (Monday May 27, 2013, was Memorial Day and Monday, September 2, 2013, was Labor Day). I split the trading day into 4 sessions based on the reported trade time: 8:00 AM – 10:59 AM (Morning), 11:00 – 1:59 PM (Mid-Day), 2:00 PM – 4:59 PM (Afternoon) and 5:00 PM – 7:59 AM (After Hours). The midday trading session was the most active, but all on-hour trading sessions have similar level of activity. About 16% of contracts were traded during the off-hour trading session. The median notional value of the contract was $50M (with a range between $1,000 and $260M). The median tenor was about 7 years (with a range between 2 months and 43 years).

Table 5 Selected contarct characteristics.

The dataset is filtered to only include observations that use USD LIBOR or CAD CDOR as the reference rate, are not voluntarily cleared, are not zero-coupon swaps and are within 50-bps of "fair rate" for the swap reported by Bloomberg. (Source: Bloomberg L.P. and author's own calculation)

	Trading Day
Monday	4,096
Tuesday	5,372
Wednesday	6,733
Thursday	6,001
Friday	5,008

	Trading Session
Morning	7,193
Mid-Day	7,845
Afternoon	7,684
After Hours	4,488

	Capped
Capped	18,837
Not Capped	8,373

	Tenor
Min	2 months
1st Quartile	5 years
Median	7 years
3rd Quartile	10 years
Max	43 years
Mean	9 years, 9 months

	Notional
Min	1,000
1st Quartile	16,000,000
Median	50,000,000
3rd Quartile	100,000,000
Max	260,000,000
Mean	56,426,143

3.6 Limitations

There are several limitations to the DTCC SDR dataset. Firstly, ~~the dataset~~ does not identify the counterparties. The identity of the counterparty (and more importantly, its creditworthiness) could have a significant impact on the swap price. In addition, the dataset does not mark which counterparty is the dealer (that is, whether the dealer is receiving the fixed rate or paying the fixed rate). When receiving the fixed rate (and paying the floating leg), the dealer is likely to require a premium over the fair price. When paying the fixed rate, the dealer is likely to require a discount below the fair price. I am also unable to observe non-standard contract characteristics such early termination provisions, collateral arrangements and day-count and settlement conventions. The standard version of the IR swaps contract uses the International Swaps and Derivatives Association (ISDA) Master Agreement for specifying these contract terms. Deviations from the ISDA Master Agreement could affect the liquidity of the contract.

4 Identification Strategy

This essay aims to analyze the impact of a policy intervention (specifically, a change in central clearing rules) on various aspects (such as prices, bid-ask spreads, and realized volatility) of certain observational units (e.g., contracts, trading days). Adapting the notation of Rubin (1976), let $Y_i(1)$ denote an outcome of interest for unit i when treated, and $Y_i(0)$ denote the outcome when untreated. The causal effect on unit i is represented by $Y_i(1) - Y_i(0)$. Since both outcomes cannot be observed simultaneously, I focus on estimating the Average Treatment Effect on the treated (ATT), $\mathbb{E}[Y_i(1) - Y_i(0) | Treatment_i = 1]$, by identifying an appropriate comparison group to estimate the counterfactual $Y_i(0)$.

Modern causal inference offers several methods to estimate this ATT, including randomized control trials (RCT), natural experiments, regression discontinuity designs (RDD), instrumental variables (IV), matching, and difference-in-differences (diff-in-diff) (Angrist and Pischke 2009; 2010; Cunningham 2021). I discuss the challenges of applying some of these methods to the research question.

In an RDD, units above a threshold value of a covariate receive treatment, while those below do not (e.g., scholarships granted to students above a specific test score). If units cannot precisely control their position relative to the threshold, assignment is “as good as random” close to the threshold, allowing for causal inference (Lee and Lemieux 2010). However, in the context of the clearing mandate, the contract characteristics (e.g., currency, notional value, floating rate index, tenor) can be precisely controlled by market participants. Thus, the assumptions necessary for RDD are not met.

The IV approach relies on an instrument that affects the likelihood of treatment assignment. When the treatment variable is endogenous, a straightforward comparison between treated and untreated units could yield biased results. For a valid IV approach, the instrument must be relevant (associated with treatment assignment), independent (free from unobserved confounding variables), and satisfy the exclusion restriction (influencing the outcome solely through treatment) (Angrist and Pischke 2009). This approach is also unsuitable here, as CFTC-defined criteria determine clearing, and no external instrument influences clearing likelihood.

The approach adopted in this essay is the diff-in-diff method. In this method, an appropriate comparison group (e.g., Canadian Dollar-denominated contracts) is selected, though it may differ from the treatment group in important ways (especially as it relates to the outcome variable). Initially, a pre-treatment difference between the comparison and treatment groups is calculated. This difference is then compared to the difference after treatment to estimate the causal effect. A key assumption in diff-in-diff is that of parallel trends—that, in the absence of treatment, both groups would have followed similar trends and the gap between the outcomes would remain constant (Angrist and Pischke 2009). This is generally not true for Canadian and U.S. swaps markets, given Canada's export-oriented economy, distinct market participants, and unique monetary and fiscal policies, all of which influence pricing, volatility, and liquidity.

However, I argue that for short periods (e.g., the 20 trading-day windows studied in this essay), the U.S. and Canadian swaps markets are highly coupled. This is supported by evidence of parallel pre-treatment trends (verified visually and through statistical methods) in pricing, volatility, and liquidity over short periods, as well as formal tests for parallel trends (see sections 4.1, 4.2, and 4.3).

4.1 Pricing

I investigate the causal impact of the central clearing mandate on the IR swap prices by comparing the premium over the fair rate (that is, the difference between the "riskless fixed rate" described in section 2.1 and the observed fixed rate on an actual contract) on USD denominated swaps versus the premium on CAD denominated swaps before and after the mandate. I employ a diff-in-diff identification strategy, with the CAD denominated swaps acting as the comparison group, which allows me to plausibly isolate the causal effect of the mandate on the swap premiums by exploiting the variation in the timing of the policy implementation.

I begin by selecting a sample of IR swaps denominated in both USD and CAD from the ten trading days before and after the central clearing mandate was implemented. I create two groups based on the currency of denomination: (1) the treatment group, consisting of USD denominated swaps that were affected by the central clearing mandate, and (2) the comparison group, consisting of CAD denominated swaps that were not subject to the mandate during the same period. By comparing the swap premiums between these two groups before and after the mandate, I can plausibly identify the causal effect of the policy on swap premiums if both groups would have followed parallel trends in the absence of the clearing mandate.

To estimate the causal effect of the central clearing mandate on swap premiums, I employ a diff-in-diff regression model, which takes the following form:

$$Y_{i,t} = \alpha + \beta_1 * Treatment_i + \beta_2 Post_t + \delta(Treatment_i \times Post_t) + X_i'\Gamma + \epsilon_{it} \tag{43}$$

where Y_{it} is the swap premium for swap i at time t, $Treatment_i$ is an indicator variable equal to 1 if the swap is denominated in USD (treatment group) and 0 otherwise (comparison group), $Post_t$ is an indicator variable equal to 1 for the period after the mandate was implemented, and $X_{i,t}$ is a vector of control variables. The coefficient of interest is δ, which captures the causal effect of the central clearing mandate on swap premiums.

The control variables included in X_i are the day of the week in which the contract is traded, the time (categorized as morning, mid-day, afternoon or off-hours) at which the contract is traded, the logarithm of the notional amount, whether the variable rate is capped, and the tenor of the contract (measured in months). The trading day variable is included because there is some discussion in the asset pricing literature of pricing differences on certain days (e.g., the "Monday effect" for equities. See (Cross 1973; French 1980)). Trading time is also similarly included to account for differences in pricing behavior during certain trading sessions during the day. For example, trading is often concentrated to the midday session, with less trading activity happening in off-hours sessions (see section 3). The lack of liquidity during those sessions can affect pricing. The logarithm of the notional value is included, as there is some discussion in the literature (Fock 2024; Randall 2015) that market participants prefer larger contracts (perhaps to economize over fixed costs) and due to the higher liquidity of these larger contracts, they may be priced differently than smaller contracts. If the notional amount is “capped,” the true nominal value of the contract is not reported. These are usually for large block trades and often involve large counterparties (“SDR View, Capped Notional Changes” 2013). Finally, the tenor (length of the contract) is included as some tenors (e.g., 10-year contracts) have more liquidity than others.

To ensure the validity of the identification strategy, I test the parallel trends assumption by both visually inspecting and statistically verifying the pre-treatment trends of swap premiums for both treatment and comparison groups and conducting placebo tests. I begin by formally testing the parallel trends assumption by regressing the fixed rate against the treatment indicator (i.e. the currency of the contract), time (trade date) and $treatment \times time$ interaction effect, for the two-week period prior to the period of study for phase 1 (that is, from Jan 28, 2013 to Feb 22, 2012). I also include some controls for contract characteristics (a subset of the control variables discussed earlier). I run the regression:

$$Y_i = \alpha + \beta_1 * Treatment_i + \beta_2 Time + \delta(Treatment \ \times Time) + X'_{i,t}\Gamma + \epsilon_{it}$$

where: $Treatment$ is an indicator variable of whether the contract is in the comparison group (currency is CAD) or treatment group (currency is USD) and $time$ is now a *continuous* variable (the trade date). This regression is run separately for each tenor of contract (since pricing for different tenors are different). Table 6 shows the results of such a regression for the two-year, five-year, and ten-year contracts. The interaction term (Currency: USD * `Trade Date) is not significant, suggesting the two groups were following parallel trends prior to the implementation of the mandatory clearing policy.

Table 6 Parallel Trends Pre-Trend tests.

The period of data is from contracts traded between Jan 28 – Feb 22, 2013 (i.e. ten trading days prior to the study period of phase 1). If parallel trends hold, we expect the interaction term (Currency Trade Date) to be not statistically significant. (Standard errors are in parentehsis).*

Pre-trend Analysis

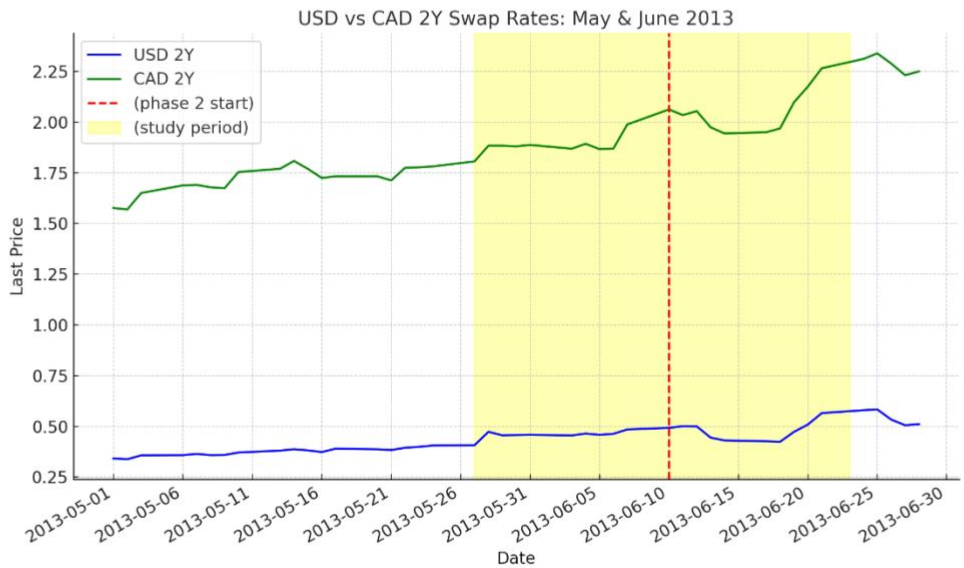
	Dependent variable:		
	Fixed Rate		
	2-year contracts	5-year contracts	10-year contracts
Currency: USD	-18.295	-58.045	-34.252
	(130.189)	(64.300)	(63.463)
Trade Date	-0.002	-0.0004	-0.0003
	(0.002)	(0.001)	(0.003)
Log Notional	0.004	-0.009	0.033^{***}
	(0.014)	(0.036)	(0.013)
Capped: Yes	0.163	0.173^{*}	0.084
	(0.161)	(0.098)	(0.052)
Currency: USD * `Trade Date	0.001	0.004	0.002
	(0.008)	(0.004)	(0.004)
Constant	37.464	8.310	6.046
	(24.307)	(20.064)	(45.747)
Observations	162	635	921
R^2	0.340	0.064	0.031
Adjusted R^2	0.319	0.057	0.026
Residual Std. Error	0.524 (df = 156)	0.740 (df = 629)	0.521 (df = 915)
F Statistic	16.071^{***} (df = 5; 156)	8.651^{***} (df = 5; 629)	5.950^{***} (df = 5; 915)
Note:		*p<0.1; **p<0.05; ***p<0.01	

I also examine whether the parallel trends assumption holds by visually inspecting the swap rate (fixed rate of the IR swaps contract) prior to each phase of the implementation of the clearing mandate. I examine the three most common tenors of USD and CAD IR swaps (2-year, 5-year and 10-year swaps). The data is reported by Bloomberg and is usually the average of 11 or more contracts traded around 11:00 AM Eastern Time of the trading day that meet contract specifications (described earlier). Figure 4 shows a sample of these trends for the tenors specified above, showing the pre-trend for phase 1 for 10-year contracts and the pre-trend for phase 2 for 2-year contracts. Other periods and phases show similar parallel trends but are not included here for brevity. The swaps rates show a parallel pre-trend prior to the implementation of the clearing mandate, with the Canadian swaps rate always higher than the US rate. This is likely due to differences in key policy rates between the US (lower Fed Funds target rate 0%) and Canada (key policy rate target 1%). There was no change to these policy rates in 2013.

Finally, I test the validity of the parallel trends assumption doing a placebo diff-in-diff. I pick the 20 trading days before the study period. I create a "placebo" diff-in-diff, as if there was a transition to clearing mandate on the 11th trading day. That is, I run the same type of diff-in-diff described earlier but pick a period when no treatment actually took place. Table 7 shows the results. The placebo diff-in-diff does not show an increase in premia (coefficient of the Group * Period term), further strengthening our belief that the increase in premia discussed in the results section (see sections 5.1 and Table 8, Table 9) is real. I forego the discussion of the control variables. They are discussed in more details in section 5.1).

Figure 4 Pre-trends for swap pricing for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation.

Red, dashed vertical line indicates when the clearing mandate went into effect. Highlighted area is the period of study and the pre-trend is to the left of highlighted area.



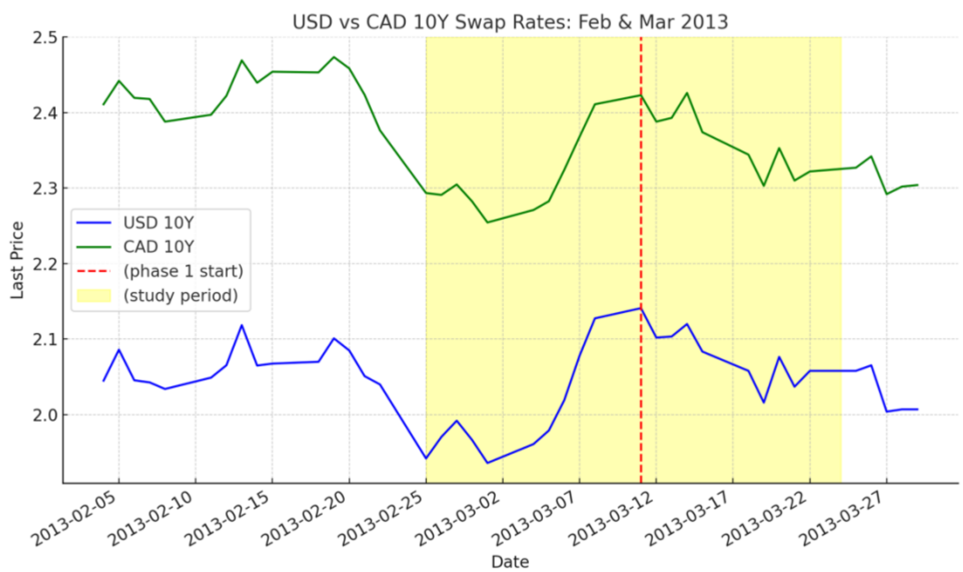


Table 7 Placebo diff-in-diff results.

*The placebo diff-in-diff is the same type of diff-in-diff analysis described in the identification strategy section, but during a period where there was no variation in central clearing requirements. I examine the 20 trading days before the study period of each phase and perform a "placebo" analysis as if the central clearing mandate had been implented on the the eleventh trading day. If the diff-in-diff method is sound, we should not see the interaction Group * Period term be significant. However, if the method shows spurious results, then Group * Period could be significant.*

Placebo Diff-in-Diff Results

	Dependent variable: Premium	
	Basic Model[6]	Advanced Model
	(1)	(2)
Group	1.6566***	1.4077***
	(0.4343)	(0.4357)
Period	-0.5706	-0.4838
	(0.5826)	(0.5799)
Tenor		0.0301***
		(0.0077)
Log Notional		-0.0219
		(0.0595)
Capped		-0.8827***
		(0.1572)
Morning Session		0.2761*
		(0.1520)
Afternoon Session		0.3836**
		(0.1600)
Off Hours		0.0617
		(0.1802)
Monday		0.7955***
		(0.1954)
Tuesday		0.5999***
		(0.1800)
Thursday		1.7587***
		(0.1744)
Friday		1.7827***
		(0.1822)

[6] The basic model regresses premia against group (USD or CAD), period (pre-treatment or post-treatment) and group×period interaction (parameter of interest). The advanced model adds additional covariates: tenor, log notional, capped, trading day and trading session (time).

Group * Period	0.1694	0.1696
	(0.5975)	(0.5952)
Constant	-1.2876***	-1.9215*
	(0.4197)	(1.1216)
Observations	20,794	20,794
R^2	0.0020	0.0136
Adjusted R^2	0.0019	0.0130
Residual Std. Error	8.4872 (df = 20790)	8.4398 (df = 20780)
F Statistic	13.8615*** (df = 3; 20790)	22.0214*** (df = 13; 20780)
Note:		*p<0.1; **p<0.05; ***p<0.01

4.2 Liquidity

Liquidity is a broad concept defined as how easily one can convert financial assets to cash with minimal impact to prices of those financial assets. Several measures try to capture this notion: *bid-ask spreads* measure the cost of trading (small quantities) of an asset at the best quoted sell and buy prices; *market depth* measures how much of an asset is available to trade at given price points; *price impact* measures how much prices react to buying and selling. As in the case of pricing (premium), I employ a diff-in-diff strategy to identify the impact of the central clearing mandate on liquidity, with CAD denominated contracts serving as the comparison group and USD denominated contracts serving as the treatment group. I use several measures of liquidity to capture the different notions discussed above.

As discussed in section 2.2 risk-averse dealers adjust their bid-ask spreads in the face of market orders from liquidity traders. To be able to compare spreads for contracts with different prices, I calculate a relative bid-ask spread based on the mid-quote:

$$spread = \frac{bid - ask}{\frac{bid + ask}{2}} \tag{44}$$

where the numerator $(bid - ask)$ is the "raw" spread and the denominator $\frac{bid+ask}{2}$ is the mid-quote. I collect end-of day bid and ask data from the Bloomberg terminal for US-dollar and Canadian-dollar denominated 2-year, 5-year and 10-year contracts for the ten-day trading period before and after the implementation of phase 1, phase 2 and phase 3 of the of the clearing regulation.

The bid-ask spread measure is "low frequency" (collected at the end of day) and does not necessarily reflect the costs traders may face throughout the trading session. To estimate the intraday liquidity that traders may experience, I use a proxy common in the literature: the Roll measure. The measure, proposed by Roll (1984) estimates the effective spread using the time series of observed prices under assumptions of market efficiency:

$$Roll = 2 \cdot \sqrt{-cov(\Delta p)} \tag{45}$$

where $cov(\Delta p)$ is the first-order serial covariance of price changes. I group the contracts by tenor and calculate the Roll measure for each trade date. I focus the analysis on 2-year, 5-year and 10-

year contracts (the most actively traded contracts). For some trading days and tenors, the quantity under the radical is negative. These observations are excluded from the dataset. I also exclude any days where the number of contracts traded is less than five.

~~In general, market liquidity is a measure of "how easily" traders can exit or enter the market.~~ The bid-ask spread ~~(~~and the Roll measure~~)~~ is a common metric for how easily traders can enter or exit *small positions*. However, when trading in larger positions, this liquidity may not be available (only a limited number of contracts may be available to trade at the best bid and ask prices, and dealers can change their quoted prices in the face of large market orders). A common measure of the "price impact" of a trade is the Amihud illiquidity measure:

$$Amihud_t = \frac{|R_t|}{Volume_t} = \frac{|P_{t-1} - P_t|}{Volume_t} \tag{46}$$

where R_t is the return on an asset at time t and $Volume_t$ is the total volume of contracts traded (calculated as the sum of the gross notional contract value in period t). The Amihud illiquidity measure, which normalizes inter-period price changes by the market size (volume) is a measure of how much prices move, scaled by the trade size. I calculate a daily Amihud illiquidity measure for US and Canadian 2-year, 5-year, and 10-year contracts for each trading date (using the contract notional as the swap equivalent of the volume and using the 1-period difference in the fixed rate of contracts with same characteristics as the return on the asset).

For each liquidity measure discussed above, I run a diff-in-diff model:

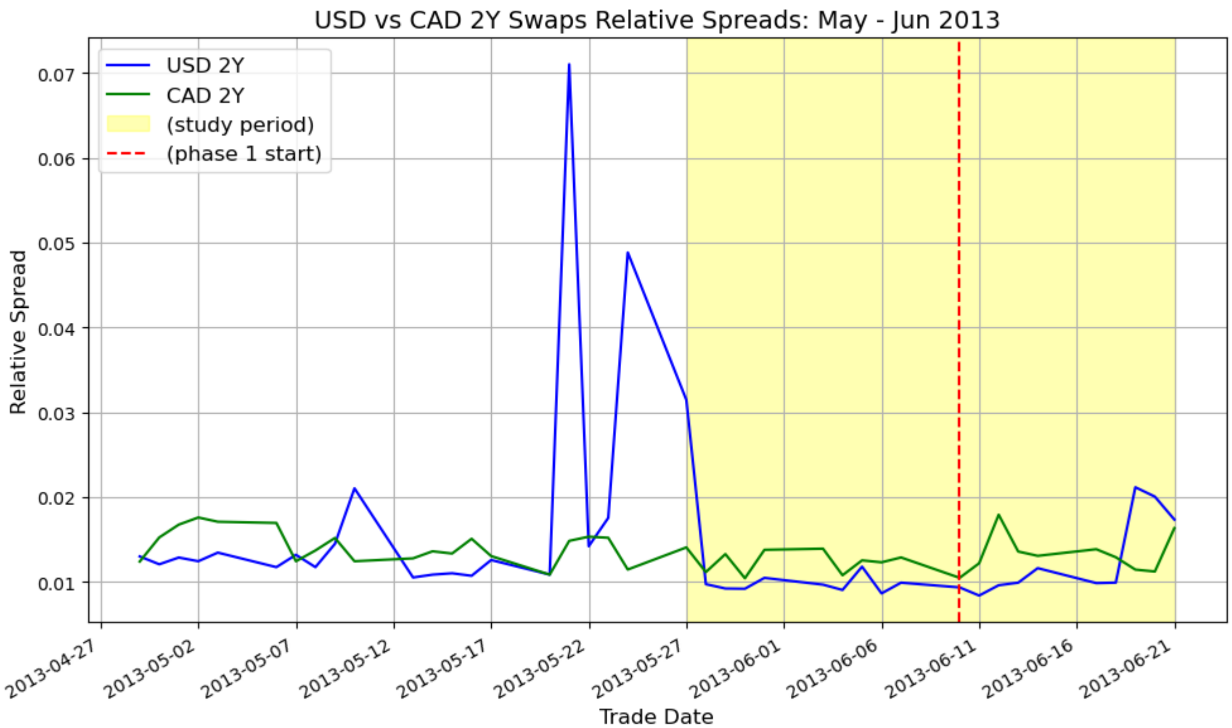
$$Y_{i,t} = \beta_0 + \beta_1 Treatment_i + \beta_2 Post_t + \delta(Treatment_i \times Post_t) + \Gamma X_{i,t} \epsilon_{i,t} \tag{47}$$

where $Y_{i,t}$ is the liquidity measure of interest, $Treatment_i$ is an indicator variable for whether the observation is in the treatment (US-dollar denominated market) or comparison (Canadian-dollar denominated market) group, $Post_t$ is an indicator variable for whether the observation is in the post- or pre implementation period. δ is the parameter of interest and $X_{i,t}$ are control variables (return on equities and equity market volatility).

For the identification strategy to be valid, the underlying variables must follow parallel trends if there was no intervention. I visually test this by plotting the time-series of the relative spread, Amihud illiquidity measure, and Roll measure for twenty trading days before the period of the study. Figure *5* shows the path of the relative spread measure for 2-year contracts and 10-year contracts in phase 1 and phase 2 respectively (other contract tenors show a similar trend). As in the case for the price premium, the relative bid-ask spread measure in the US and Canadian contracts generally follow a parallel trend before the implementation of the clearing mandate. The Canadian market has higher relative spreads, likely due to the smaller size of the market. Figure *6* shows the Roll measure for the 2-year and 10-year USD and CAD IR swaps contracts for phase 1 and 2 respectively (other tenors show similar trends). Note that due to the limitation previously discussed (negative covariance or days with limited trading volume), the Roll measure is not observable on all trading days. It is difficult to make a conclusion about parallel trends due to the lack of data availability. Figure 7 shows the Amihud illiquidity measure again for 2-year and 10-year USD and CAD IR swaps contracts in phase 1 and phase 2, respectively. As expected, the Amihud illiquidity measure is larger (indicating less liquidity) for the Canadian market, due to its smaller size. As is the case with the Roll measure, the Amihud illiquidity measure is not computable for each tenor for each day (for example, if there are less than 5 observations for a given contract specification on a given day). There is not enough data to conclude whether the parallel trend assumptions hold for the Amihud illiquidity measure.

Figure 5 Pre-trends for Relative bid-ask spreads for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation.

Red, dashed vertical line indicates when the clearing mandate went into effect. Highlighted area is the period of study and the pre-trend is to the left of highlighted area.



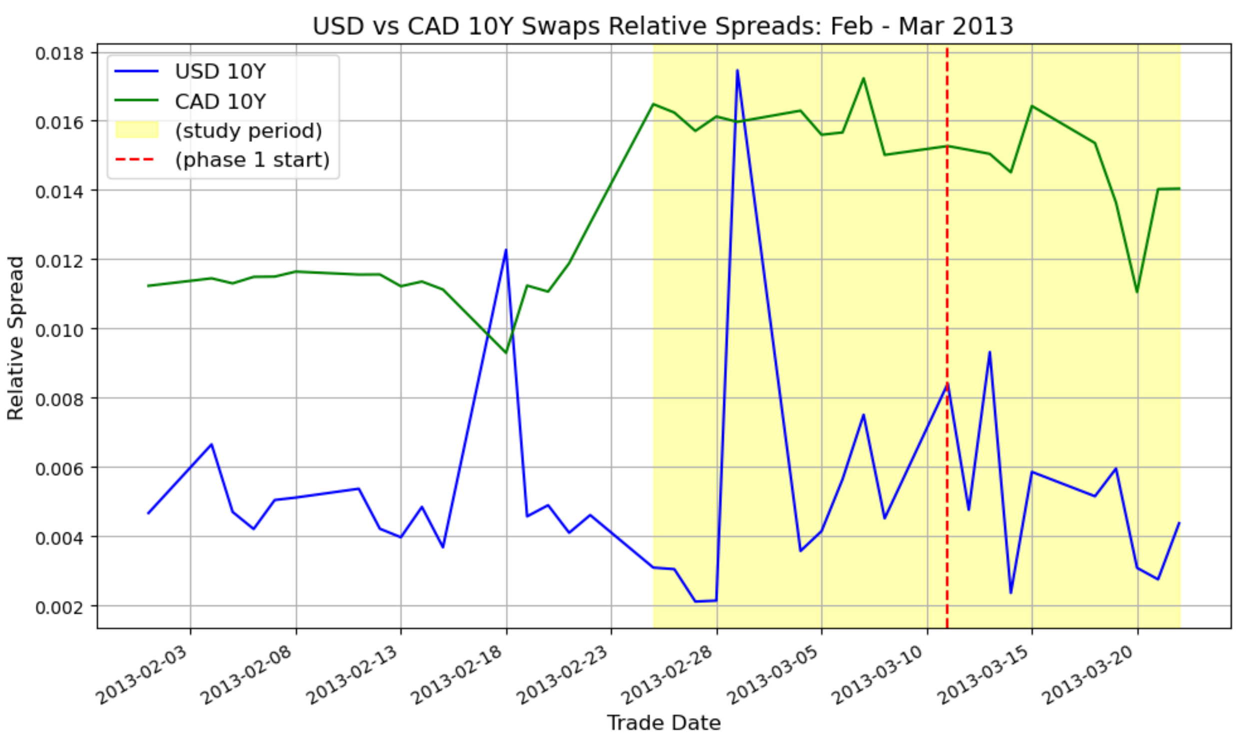
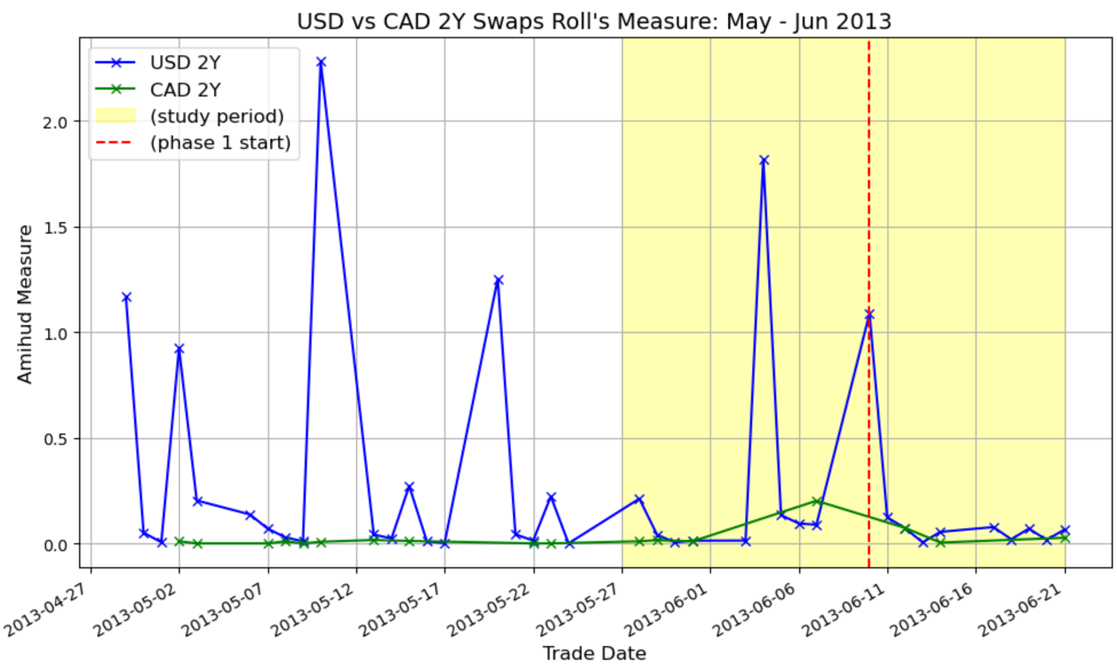


Figure 6 Pre-trends for Roll's Measure for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation.

Red, dashed vertical line indicates when the clearing mandate went into effect. Highlighted area is the period of study and the pre-trend is to the left of highlighted area. Markers (x) show the value at a given date. Data are not available for all contracts at all dates.



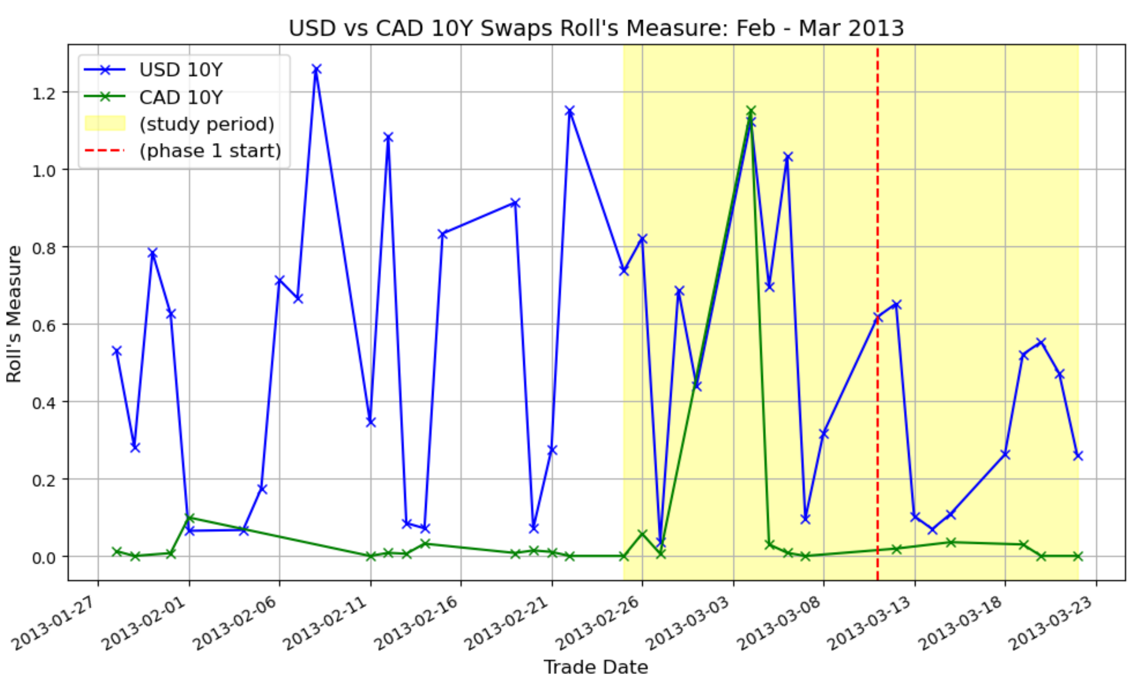
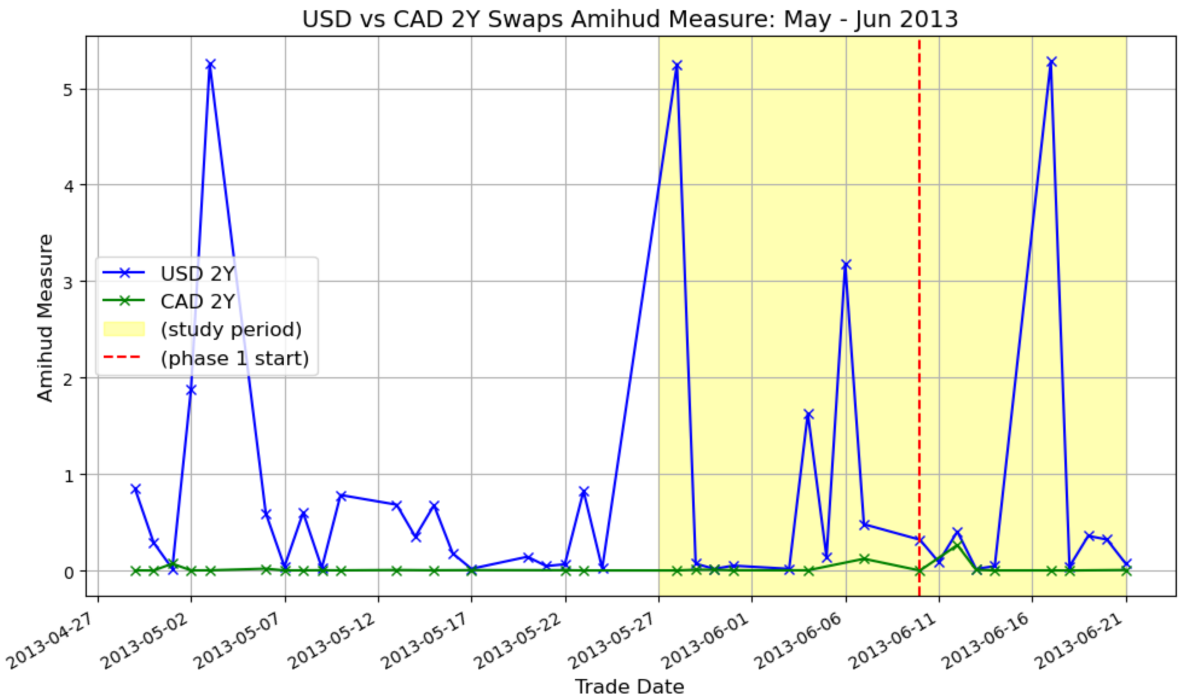
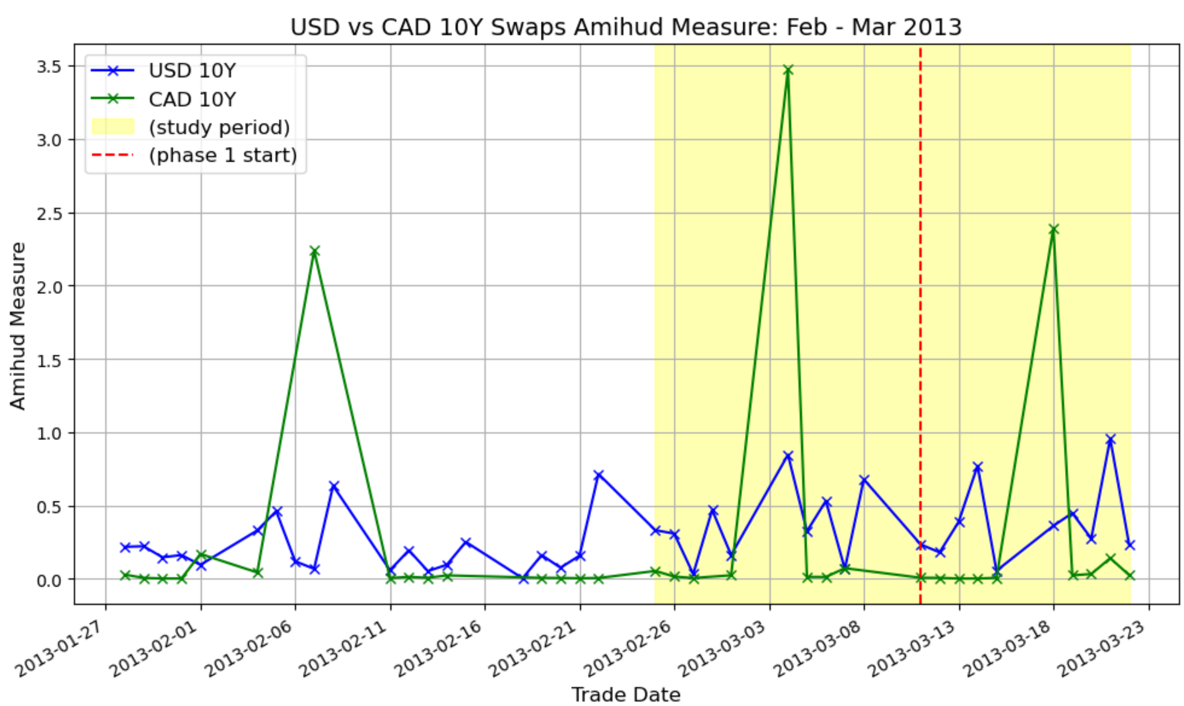


Figure 7 Pre-trends for Amihud Illiquidty Measure for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation.

Red, dashed vertical line indicates when the clearing mandate went into effect. Highlighted area is the period of study and the pre-trend is to the left of highlighted area. Markers 'x' show the value at a given date. Data are not available for all contracts at all dates.





4.3 Volatility

As is widespread practice in literature, I use the realized volatility as my measure of volatility. If the return on an asset in the period $[t-1, t]$ is defined as:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{48}$$

The (annualized) realized volatility of the return is then:

$$RV_t = \sum 252 \cdot (r_t)^2 \tag{49}$$

where: n is the number of trading days in the sample period and 252 is the approximate number of trading days in a year.

For each trading day, I select contracts with "whole number tenor years" between 1 and 10 years, as well as 15-years and 30-years (I exclude contracts that are "partial years" such as 18-, 21- and 30-month contracts). For the Canadian market, these are the most actively traded contracts. Calculating volatility requires several observations of each tenor for each trading day. I group contracts by currency, tenor, and trading day (I exclude observations when less than 5 contracts with a given specification are traded on a given day). The filtered dataset captures 90% of Canadian contracts traded during the period studied in the essay, and I can calculate volatility of several tenors for each trading day[7]. However, for some tenors (such as 4-year, 6-year, 8-year and 9-year contracts), no trades or only one or two trades occur in the Canadian market on certain dates, and I cannot calculate the volatility measure for that tenor for the Canadian-dollar denominated contract on that trading date. Ideally, I would have 24 observations for each of the 60 trading days (1,392 observations in total, one for each currency-tenor combination for each day). However, since sometimes no Canadian contract of a particular tenor is traded on certain dates, I end up with 914 observations in my data set.

Like liquidity and pricing calculations, I classify each observation as either in the comparison group (if currency is CAD) or treatment group (if currency is USD), and whether it is in the pre-treatment or post-treatment period. I then perform a diff-in-diff regression:

$$RV_{i,t} = \beta_0 + \beta_1 Treatment_i + \beta_2 Period_t + \delta(Treatment_i \times Period_t) + X'_{i,t}\Gamma + \epsilon_{i,t} \tag{50}$$

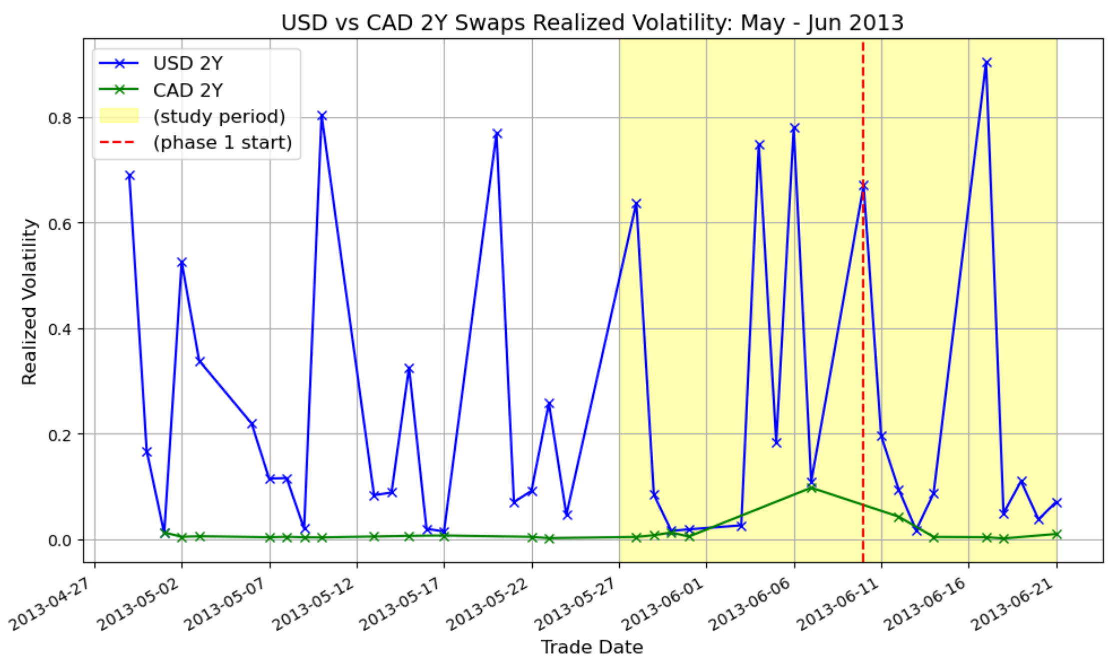
[7] I exclude discussion regarding the data availability of US-dollar denominated contracts. In general, there are enough contracts traded for each trading day and contract specification that a realized volatility measure can be calculated for each. The availability of data for the US market is not a limiting factor in my analysis.

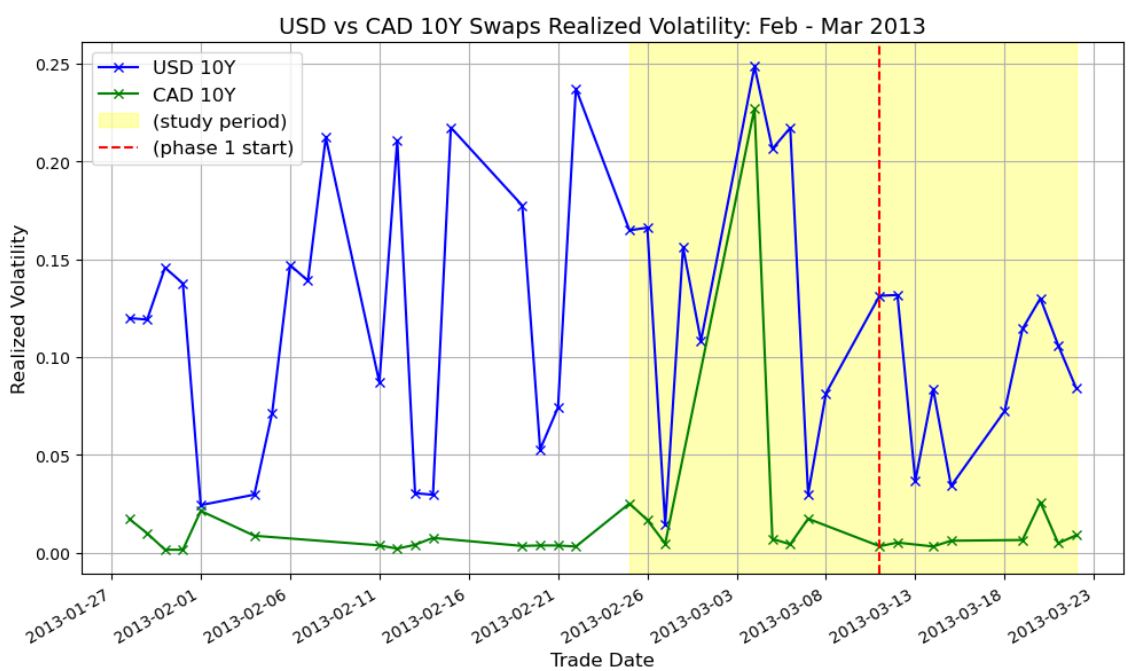
where $RV_{i,t}$ is the realized volatility for contract specification i in period t and the rest of the variables are as described in the liquidity section. δ (the interaction between group and pre/post-treatment period) is the parameter of interest. I use the same controls (equity market returns and equity market volatility) as in the liquidity section.

As with the other diff-in-diff specification, for the identification strategy to be valid, the two groups need to follow parallel trends in the absence of an intervention. I plot the time series (Figure 8) of the realized volatility measure for the 2-year, and 10-year contracts for the twenty trading days before the implementation of the clearing mandate. In general, the volatility measure follows a parallel trend for the US-dollar denominated and Canadian-dollar denominated markets.

Figure 8 Pre-trends for Realized Volatility for 2-year swaps during phase 2 and 10-year swaps during phase 1 of the clearing mandate implementation.

Red, dashed vertical line indicates when the clearing mandate went into effect. Highlighted area is the period of study and the pre-trend is to the left of highlighted area. Markers 'x' show the value at a given date. Data are not available for all contracts at all dates.





5 Results

5.1 Pricing

For analyzing the impact of the clearing mandate on prices, I compare USD denominated contracts using LIBOR as the floating rate index, against CAD denominated contracts using the CDOR as the floating rate index. The USD LIBOR contracts are subject to the CFTC clearing mandate. Table 8 lists the diff-in-diff results for the swap premium, pooling data from all phases. Column 1 shows a basic model without any controls for contract characteristics. Column 2 (advanced model) shows the effects additional controls, such as the (log) notional value of the contract, day, and period of trading and whether the notional value was "capped" (i.e., the exact value was not reported to the trade repo).

The clearing mandate causes a ~14 bps rise in premia across the three phases in this model per the basic model. In the more advanced model with additional controls for contract characteristics, premia rise by ~13 bps. These results are qualitatively in line with the theoretical model that reducing the riskiness of the contract increases its price.

Examining the control variables, beginning with the trading day, and using Wednesday as the reference level, I note that there is a 1.0-3.0 bps increase in the premium depending on the trading day. There is also a 1.0-1.3 bps decrease in the premium for trading in morning, afternoon or off hours trading sessions (as compared to midday). Both results contrast with assumptions of "efficient markets," where there should be no arbitrage opportunities by trading during special days or times. A one percent increase in the notional value is associated with a 0.77 bps increase in the premium. Again, this contrasts with expectations from "efficient market" assumptions because arbitrage opportunities exist (for example, a dealer can make a riskless profit by agreeing to receive a fixed rate on a higher-priced "large" contract and agreeing to pay the fixed-rate for two lower-priced "small" contracts). Finally, a one-year increase in the tenor is associated with a 0.03 bps increase in the premium. It is difficult to determine whether this represents a riskless-arbitrage opportunity or a difference in perceived riskiness for longer-dated contracts. Although many of the covariates are statistically significant, the magnitudes of the effects are small, ranging from 0.03 to 3 bps.

Table 9 shows the result of a diff-in-diff analysis on each phase separately. In phase 1, there is a ~5.3 bps increase in premia after the implementation of the mandate. As noted previously, there was a 16% increase in the cleared volume following implementation of phase 1. In phase 2, there was an additional ~2.6 bps increase in premia. In phase 3, premia increased by ~16 bps. These results generally hold to the idea that as more of the market is cleared, there is a perceived reduction in counterparty risk and swap premia rise. The results are consistent with the pooled diff-in-diff, with most of the effect occurring during the third phase of the mandate.

Table 10 shows the results of a similar regression using an alternative currency pair. The CFTC clearing mandate also affected contracts denominated in GBP using the LIBOR as the reference rate (with the same implementation dates as the USD clearing mandate). These contracts now serve as the treatment group. The clearing mandate did not apply to Swiss Franc (CHF) denominated contracts, and these contracts now serve as the comparison group. The clearing mandate had a similar (but smaller) impact on prices of GBP-denominated swaps, further strengthening our belief that clearing reduces counterparty risk and increases contract premia.

Table 8 diff-in-diff results for prices pooling all phases together.

*Column 1 shows results for a basic diff-in-diff model without controlling for any covariates. Column 2 cotrols for contract charatceristics such as the tenor (in years), (log) notional value, whether notional is capped, whether the swap was traded on an eelctronic swap execution facility, the trading session and the trading day. The parameter of interest is the interaction term Group * Period, which shows an effect of 13.4-14.2 bps increase in premia for the treatment group once clearing is enacted.*

Diff-in-diff Regression Results

	Dependent variable: Premium	
	Basic Model	Full Model
	(1)	(2)
Group	-0.8889^{*}	-0.7683
	(0.4917)	(0.4900)
Period	-13.6369^{***}	-13.2955^{***}
	(0.6641)	(0.6610)
Tenor		0.0362^{***}
		(0.0086)
Log Notional		0.7755^{***}
		(0.0671)
Capped		-0.9311^{***}
		(0.1849)
SEF		0.6922
		(2.5197)
Morning Session		-1.0238^{***}
		(0.1843)
Afternoon Session		-1.2368^{***}
		(0.1814)
Off Hours		-1.2907^{***}
		(0.2125)

Monday		1.5672^{***}
		(0.2244)
Tuesday		2.3944^{***}
		(0.2070)
Thursday		2.7672^{***}
		(0.2005)
Friday		0.9566^{***}
		(0.2124)
Group * Period	14.2183^{***}	13.4103^{***}
	(0.6833)	(0.6839)
Constant	-0.2415	-14.1707^{***}
	(0.4718)	(1.2407)
Observations	27,210	27,210
R^2	0.0283	0.0444
Adjusted R^2	0.0282	0.0440
Residual Std. Error	11.3530 (df = 27206)	11.2607 (df = 27195)
F Statistic	264.3342^{***} (df = 3; 27206)	90.3482^{***} (df = 14; 27195)
Note:		*p**p***p<0.01

Table 9 diff-in-doff results for prices by each implementation phase separately.

*This analysis uses the same "full model" described earlier (controlling for contract charatceristics). There is a 5.3 bps increase in premia in phase 1, a 2.7 bps increase in premia in phase 2 and a 16.2 bps increase in premia in phase 3 (coefficients on the Gorup * Period term).*

By Phase Results: Advanced Model

	Dependent variable: Premium		
	Phase 1	Phase 2	Phase 3
	(1)	(2)	(3)
Group	-2.789***	2.327***	3.139***
	(0.525)	(0.886)	(1.205)
Period	-4.898***	-4.150***	-12.360***
	(0.875)	(1.309)	(1.338)
Tenor	-0.050***	0.064***	0.086***
	(0.013)	(0.013)	(0.016)
Notional	-0.489***	0.685***	1.506***
	(0.094)	(0.109)	(0.125)
Capped	-0.727***	-0.583**	-1.575***
	(0.268)	(0.287)	(0.345)
Morning Session	-0.387	0.788***	-2.375***
	(0.265)	(0.292)	(0.340)
Afternoon Session	-1.170***	-0.571**	-0.538
	(0.264)	(0.280)	(0.342)
Off Hours	-1.196***	1.594***	-5.542***
	(0.309)	(0.334)	(0.392)
Monday	2.017***	6.666***	-5.821***
	(0.323)	(0.367)	(0.409)
Tuesday	0.741**	8.913***	-3.854***

	(0.312)	(0.326)	(0.377)
Thursday	2.025***	8.909***	-3.700***
	(0.306)	(0.306)	(0.376)
Friday	1.642***	5.832***	-4.480***
	(0.325)	(0.320)	(0.402)
Group * Period	5.308***	2.658**	16.277***
	(0.899)	(1.336)	(1.408)
Constant	11.804***	-22.064***	-27.840***
	(1.654)	(2.101)	(2.446)
Observations	7,561	10,856	8,793
R^2	0.025	0.109	0.179
Adjusted R^2	0.024	0.108	0.178
Residual Std. Error	8.635 (df = 7547)	11.002 (df = 10842)	11.861 (df = 8779)
F Statistic	15.068*** (df = 13; 7547)	102.336*** (df = 13; 10842)	147.232*** (df = 13; 8779)
Note:			*p**p***p<0.01

Table 10 diff-in-diff using alternative Currency Pair

*GBP denominated contracts serve as the treatment group and CHF denominated contracts serve as the comparison group. The interaction term Group * Period shows a 7.5-8.3 bps increase (depending on model) for the treatment group once clearing becomes mandatory.*

Alternative Currencies Diff-in-diff Results (GBP vs. CHF)

	Dependent variable: Premium	
	Basic Model	Advanced Model
	(1)	(2)
Group	-0.2734	-0.9492
	(1.2232)	(1.2023)
Period	-6.6242^{***}	-8.2303^{***}
	(1.4576)	(1.4435)
Tenor		0.0974^{***}
		(0.0193)
Log Notional		0.6572^{***}
		(0.1811)
Capped		-0.4361
		(0.5052)
Morning Session		-0.9820^{**}
		(0.4967)
Afternoon Session		-2.5981^{***}
		(0.4186)
Off Hours		-2.4152^{***}
		(0.7555)
Monday		3.1430^{***}
		(0.5643)
Tuesday		3.5697^{***}
		(0.5050)
Thursday		3.0135^{***}
		(0.4984)
Friday		1.5464^{***}
		(0.5515)
Group * Period	7.4610^{***}	8.2859^{***}
	(1.5404)	(1.5143)
Constant	-3.7350^{***}	-15.0964^{***}

	(1.1343)	(3.1718)
Observations	3,522	3,522
R^2	0.0168	0.0580
Adjusted R^2	0.0159	0.0546
Residual Std. Error	10.3965 (df = 3518)	10.1905 (df = 3508)
F Statistic	20.0170*** (df = 3; 3518)	16.6288*** (df = 13; 3508)
Note:		*p<0.1; **p<0.05; ***p<0.01

5.2 Liquidity

As noted previously, liquidity is a broad concept with several measures. I begin by examining the impact of the central clearing mandate on the relative bid-ask spread prevalent at the end of the trading day, a measure of the trading cost ~~(at least at the closing)~~ scaled for the price of the contract. Table 11 shows the results of a diff-in-diff regression for this measure. In the full model, I include the contract tenor as a control variable, as the relative bid-ask spread varies significantly with the contract tenor (for example, the median relative bid-ask spread in the dataset for Canadian 2-year contract was 0.015% while for a 10-year contract was 0.012%. For US contracts, the median relative bid-ask spread was 0.01% for 2-year contracts and 0.003% for the much more heavily traded 10-year contracts). Since the period of study is relatively short and since liquidity is a "market wide," rather than an individual contract-based measure, the opportunity to control for variables that impact liquidity is limited to market-wide metrics[8]. If a longer period were being studied, variables that impact liquidity, such as monetary policy and credit availability could be added as controls. However, these variables did not vary during the brief period studied and cannot be controlled for. Two control variables that proxy financial market conditions are added to the full model: a measure of equity market volatility and a measure for equity market return. For the volatility measure, I use the CBOE Volatility Index (CBOE VIX) and the TSX 60 VIX Index, which measure the 30-day expected realized variance of the S&P 500 Index and its Canadian equivalent, respectively. The Canadian VIX index was launched in April 2021, but S&P provides hypothetical historical values. For equity market returns, I use the returns of the S&P 500 and the S&P/TSX Composite.

The diff-in-diff results suggest that the clearing mandate does not impact liquidity as measured by relative bid-ask spreads. In the theory section, I argued that we should expect reductions in counterparty risk to cause a narrowing of the bid-ask spread, as the spread is charged by dealers to offset their expected losses from holding inventory, and a reduction in counterparty risk reduces these expected losses. However, the spread is also driven by supply and demand conditions in the market (i.e., the quantity and market order size discussed in the theory section). A reduction in

[8] That is, it doesn't make much sense to talk about the liquidity of a particular 10-year contract. Rather, we look at the liquidity for all 10-year standardized swaps contracts on a given trade date.

riskiness of IR swaps increases their demand. If the swaps market is monopolistic (that is, new swaps dealers face barriers to entry), then incumbent dealers can choose not to adjust their bid-ask spreads and pocket the additional profits from the higher demand.

I examine two additional measures of liquidity. Roll's measure is an estimate of bid-ask spreads that might have prevailed during the trading day and is obtained from transaction data (price and trade time). ~~The relative bid-ask spread above is based only on the last quote and is likely indicative of liquidity costs at the end of the trading day.~~ Table 12 shows the results of a diff-in-diff regression for this measure. The contract tenor as well as the equity market volatility and equity market return variables from the previous discussion are included as control variables in the full model. ~~As the~~ relative bid-ask spread, Roll's measure does not show a statistically significant change in trading costs due to the clearing mandate.

The Amihud illiquidity measure is an estimate of the average price impact ~~(by what percentage prices change for a given order of size).~~ I use the notional contract amount as the "order size" and the (log) difference in the fixed rate between two trades as the (percent) change in price. I express the results in percent change per million dollars of order quantity for easier interpretation. Table 13 show the results for the Amihud illiquidity measure diff-in-diff analysis. The Amihud illiquidity measure shows a statistically significant but small (0.36% change in price/million dollars) impact of the clearing mandate.

Table 11 diff-in-diff analysis of relative bid-ask spreads.

The dependent variable is the Relative Bid-Ask Spread from the last quote of the trading day. For the full model, control variables are the contract tenor, relevant stock market (TSX or S&P 500) returns and volatility. Stock market returns are calculated as the percent change from the previous trading day's adjusted closing price (where adjustments are made for dividends, stock splits and other rights offers). The volatility measures are the CBOE VIX and TSX 60 VIXI indices.

Relative Bid-Ask Spread diff-in-diff Analysis

	Dependent variable:	
	Relative Spread	
	Simple Model	Full Model
	(1)	(2)
Group	-0.005^{***}	-0.005^{***}
	(0.001)	(0.001)
Period	0.0001	-0.0001
	(0.001)	(0.001)
Tenor (2Y)		0.006^{***}
		(0.001)
Tenor (5Y)		0.003^{***}
		(0.001)
Equity Return		-0.056^{*}
		(0.031)
Volatility		-0.0003^{**}
		(0.0001)
Group*Period	-0.001	-0.001
	(0.001)	(0.001)
Constant	0.014^{***}	0.016^{***}
	(0.001)	(0.002)

Observations	360	351
R^2	0.204	0.486
Adjusted R^2	0.198	0.476
Residual Std. Error	0.005 (df = 356)	0.004 (df = 343)
F Statistic	30.459^{***} (df = 3; 356)	46.369^{***} (df = 7; 343)
Note:		*p**p***p<0.01

Table 12 diff-in-diff analysis of Roll's Measure.

The dependent variable is the daily Roll's Measure (a proxy of the bid-ask spread during the trading day). Control variables are the contract tenor, relevant stock market (TSX or S&P 500) returns and volatility. Stock market returns are calculated as the percent change from the previous trading day's adjusted closing price (where adjustments are made for dividends, stock splits and other rights offers). The volatility measures are the CBOE VIX and TSX 60 VIXI indices.

Roll's Measure diff-in-diff Analysis

	Dependent variable:	
	Roll's Measure	
	Simple Model	Full Model
	(1)	(2)
Group	0.368^{***}	0.377^{***}
	(0.041)	(0.040)
Period	0.040	0.042
	(0.063)	(0.062)
Tenor (2Y)		-0.113^{***}
		(0.042)
Tenor (5Y)		0.044
		(0.040)
Equity Return		4.289^{*}
		(2.591)
Volatility		-0.009
		(0.011)
Group * Period	-0.139^{*}	-0.133^{*}
	(0.079)	(0.078)
Constant	0.069^{**}	0.213
	(0.032)	(0.159)

Observations	551	548
R^2	0.149	0.183
Adjusted R^2	0.144	0.173
Residual Std. Error	0.393 (df = 547)	0.387 (df = 540)
F Statistic	31.924^{***} (df = 3; 547)	17.303^{***} (df = 7; 540)
Note:		*p**p***p<0.01

Table 13 diff-in-diff analysis of Amihud's Illiquidity Measure.

The dependent variable is Amihud's Measure (expressed in absolute % change in the fixed rate of the contract per a million dollar of notional value traded). Control variables are the contract tenor, relevant stock market (TSX or S&P 500) returns and volatility. Stock market returns are calculated as the percent change from the previous trading day's adjusted closing price (where adjustments are made for dividends, stock splits and other rights offers). The volatility measures are the CBOE VIX and TSX 60 VIXI indices.

Amihud lliquidity Measure diff-in-diff Analysis

	Dependent variable:	
	Amihud Illiquidity Measure	
	Simple Model	Full Model
	(1)	(2)
Group	0.348^{***}	0.367^{***}
	(0.068)	(0.068)
Period	0.362^{***}	0.364^{***}
	(0.099)	(0.098)
Tenor (2Y)		0.018
		(0.072)
Tenor (5Y)		0.206^{***}
		(0.070)
Equity Return		12.328^{***}
		(4.469)
Volatility		0.025
		(0.019)
Group * Period	-0.365^{***}	-0.386^{***}
	(0.133)	(0.132)
Constant	0.106^{**}	-0.344

	(0.051)	(0.280)
Observations	641	635
R^2	0.048	0.077
Adjusted R^2	0.043	0.066
Residual Std. Error	0.733 (df = 637)	0.727 (df = 627)
F Statistic	10.674^{***} (df = 3; 637)	7.445^{***} (df = 7; 627)
Note:		*p**p***p<0.01

5.3 Volatility

I measure price volatility as the daily realized variance of the fixed-rate leg and estimate its response to the clearing mandate with the diff-in-diff specification in equation (50). Each observation is a tenor–currency–day cell; the sample pools the three mandate windows, giving 829 observations after filtering for data availability.

Table 14 reports the diff-in-diff results. Volatility is systematically higher for USD contracts than for CAD contracts: the "Group" coefficient is 6.7 bps (simple model) and 9.5 bps (full model) and is statistically significant at the 1 percent level. By contrast, the "Period" term is economically small (≈ 1 bps) and statistically insignificant, indicating no general drift in volatility over the twenty-day windows around each phase. Most importantly, the interaction term is indistinguishable from zero (0.2 bps in the simple model and –0.6 bps in the full model, both with s.e. ≈ 2.5 bps). Thus, clearing did not measurably dampen day-to-day rate volatility under "normal" market conditions.

Control variables behave as expected. Short-dated contracts are more volatile (the 1-year tenor adds 24 bps relative to the 10-year benchmark), while the 15- and 30-year contracts are less volatile. Equity-market returns and VIX-style measures have no incremental explanatory power once swap-specific covariates are included. The modest R^2 (about 3 percent in the simple specification and 25 percent in the full specification) is typical for daily RV regressions and reflects the near-random-walk behavior of swap rates in quiet periods.

Theory suggest that any stabilizing benefit of central clearing should be most visible when counterparty-credit concerns flare up. To test this prediction, I re-estimate the diff-in-diff over the June 27–July 13, 2015, window, bracketing the announcement and resolution of the Greek EU referendum. The sample ~~contracts~~ to 104 observations, and CAD swaps provide a noisy comparison because of their limited exposure to euro-area risk. Table 15 shows that the mandate's interaction term is –1.6 bps with a 11.9 bps standard error (again not statistically significanct). Realized volatility would have had to fall by roughly 35 bps (one-third of the sample mean) to achieve conventional significance levels in this small window.

Two factors likely explain the null findings. First, the model in Section 2.5 emphasizes order-flow shocks (σ) as the primary driver of short-horizon price changes; reducing counterparty-risk therefore moves $\alpha\sigma^2$ only at second order. Second, contemporaneous regulatory changes particularly greater post-trade transparency and the migration to SEF trading may have compressed volatility in both USD and CAD markets, biasing the diff-in-diff toward zero.

Table 14 diff-in-diff analysis of volatility of the fixed rate for USD and CAD IR swaps contracts

Daily volatility diff-in-diff Analysis

	Dependent variable:	
	Realized Volatility	
	Simple Model	Full Model
	(1)	(2)
Group	0.067^{***}	0.095^{***}
	(0.020)	(0.018)
Period	0.009	0.018
	(0.024)	(0.021)
Equity Mkt Return		0.906
		(0.704)
Equity Mkt Volatility		0.003
		(0.003)
Tenor 15Y		-0.053^{**}
		(0.024)
Tenor 1Y		0.240^{***}
		(0.025)
Tenor 2Y		0.047^{**}
		(0.021)
Tenor 30Y		-0.070^{***}
		(0.023)

Tenor 3Y		-0.038^{*}
		(0.022)
Tenor 4Y		-0.052^{**}
		(0.023)
Tenor 5Y		0.043^{**}
		(0.020)
Tenor 6Y		-0.053^{**}
		(0.024)
Tenor 7Y		-0.036
		(0.024)
Tenor 8Y		-0.067^{***}
		(0.024)
Tenor 9Y		-0.074^{***}
		(0.024)
Group * period	0.002	-0.006
	(0.027)	(0.024)
Constant	0.027	-0.029
	(0.017)	(0.047)
Observations	829	829
R^2	0.031	0.247
Adjusted R^2	0.028	0.232
Residual Std. Error	0.161 (df = 825)	0.143 (df = 812)
F Statistic	8.935^{***} (df = 3; 825)	16.612^{***} (df = 16; 812)
Note:	*p**p***p<0.01	

Table 15 diff-in-diff analysis of price volatility during second Grexit referendum

Diff-in-diff Analysis of Volatility During GREXIT

	Dependent variable:
	Volatility (Daily Return)
Period	-0.002
	(0.087)
Group	0.012
	(0.090)
Period x Group	-0.016
	(0.119)
Constant	0.039
	(0.065)
Observations	104
R^2	0.001
Adjusted R^2	-0.029
Residual Std. Error	0.300 (df = 100)
F Statistic	0.018 (df = 3; 100)
Note:	*p**p***p<0.01

6 Conclusion

This study investigates the causal impact of the central clearing mandate on the IR swaps ~~(IRS)~~ market, focusing on key outcomes such as pricing, liquidity, and volatility. Using a diff-in-diff approach, I can isolate the effects of the clearing mandate, providing a comprehensive view of its influence on market dynamics.

The findings suggest that central clearing plays a significant role in reducing counterparty risk, as evidenced by the consistent rise in swap premia following the mandate. This reflects an increased valuation for cleared contracts, indicating market participants place a higher premium on reduced risk exposure. However, the anticipated improvements in liquidity were not observed. Measures such as the bid-ask spread, Roll measure, and Amihud illiquidity measure show no substantial change in liquidity because of the clearing mandate. This suggests that in monopolistic or concentrated dealer markets, the ~~demand for cleared contracts~~ does not necessarily lead to narrower spreads or improved liquidity conditions.

Regarding price volatility, the results indicate that under normal market conditions, the mandate has little to no effect on volatility. The realized volatility measures reveal that prices generally follow a random walk during stable periods, making it difficult to detect significant changes due to the clearing requirement. However, during episodes of market stress, such as the event surrounding the second "Grexit" vote, cleared contracts experienced lower volatility compared to their uncleared counterparts, implying that central clearing may enhance stability in more turbulent times.

While the mandate has succeeded in reducing counterparty risk, its impact on liquidity and volatility appears more nuanced. The clearinghouse structure has not necessarily resulted in a more liquid market, and its effect on volatility is more pronounced during periods of financial stress rather than regular market conditions. These results are crucial for regulators and market participants, as they highlight both the strengths and limitations of central clearing in maintaining market stability.

Future research could delve deeper into the long-term effects of central clearing, particularly in crisis periods, and explore whether different market structures or alternative clearing mechanisms might enhance both liquidity and stability in the IRS market.

Essay 2: An Information Theoretic Extension to Conditional Differences-in-Differences

The difference-in-differences approach (sometimes abbreviated DiD or diff-in-diff) to causal identification has a long history, perhaps tracing back to work of physician John Snow in identifying the cause of the 1854 London Cholera outbreak (Cunningham 2021). In economics, precursors to diff-in-diff methods began to appear in the literature in the in the 1950s (for example (Lester 1946) as recounted in (Halkiewicz 2024)). The modern diff-in-diff approach was popularized by Ashenfelter (1978), with further crucial work by Ashenfelter and Card (1984), Angrist (1990; 1991), Angrist and Kruger(1991; 1999), and Imbens (1995). This development was part of the "credibility revolution" taking place in economics in the 1990s, which saw a variety of methods for identifying causal effects (also known as identification strategies), such as regression discontinuity designs (RDD), instrumental variables (IV), natural experiments and randomized control trials (RCT), becoming popular in the field (Angrist and Pischke 2010).

Each identification strategy relies on some assumptions about the underlying data generation process (DGP) and causal mechanism to make a valid causal inference. The identification issue arises from the "missing counterfactual" problem: we cannot observe the same units both receive and not receive the treatment (the intervention or policy variable whose effects we are interested in studying). We try to fill in this missing counterfactual with some stand-in. In the case of diff-in-diff, we identify an alternative group that never receives the treatment during our period of study. We then assume that the groups would have followed "parallel trends" for the outcome variable (that is, the gap between the treated and untreated groups would have remained constant) in the absence of any intervention. Thus, any change in the gap between the outcome variables after treatment can be attributed to the intervention.

For a concrete example, consider a diff-in-diff study that attempts to identify the effect of a job training program on wages (for example, as in (LaLonde 1986)). We observe the wages of those in the training program before and after they receive training. However, we cannot plausibly identify the impact of the training program on participants' wages based on this information alone, since we do not observe what their wages would have been had they not received training. After all, other variables (such as the business cycle) might have caused changes in wages, and if we simply compared the before and after training wages of this group, we would be mistakenly attributing the change in wages to job training, when another variable could have been the cause. We try to get around this missing counterfactual problem by identifying a group that never receives job training during the period of study. Let's say that before the job training program started, those not in the training program on average made $800 more a year than those in the training program (job training is often targeted to those who are unemployed, underemployed or low income). We now make the crucial assumption that this gap would have remained constant if not for the job training program. That is, although wages for both groups might have increased or declined during the period of study (e.g. due to the business cycle), the average gap in wages between the groups would have remained constant without the job training program. If we can plausibly make this claim, then any changes in the gap after training can be attributed to the program. For example, if after the program, the gap in wages between the two groups narrowed to $300, then the program can be attributed to an average $500 increase on the earnings of those who received it.

The diff-in-diff methodology requires the crucial and restrictive parallel trends assumption. This assumption is sometimes implausible. In the job training example, the comparison group might differ in terms of age, education, gender and racial composition from those receiving training. Let us assume that those in the job training program are on average, younger than those in the comparison group. If younger workers experienced higher wage growth throughout the economy during our period of study, then we could be falsely attributing some of the wage gains to the job training program when the gains are really driven by the difference in ages between the two groups. To make a valid causal inference in this situation, we might adopt the less restrictive "conditional parallel trends" assumption. That is, we could assume that the groups would have followed parallel trends conditionally on some values of the covariates.

The literature suggests several ways to estimate this type of conditional diff-in-diff model. Firstly, if we have a good model of how the outcome variables would have evolved conditional on the covariates (for example, if we know how age affects wages), we can incorporate this into the two-way fixed-effects (TWFE) regression commonly used to estimate the classical diff-in-diff model (Baker et al. 2025). Other methods include matching and inverse-probability weighting (IPW) (Abadie 2005; Abadie and Imbens 2006), regression adjustment (J. J. Heckman, Ichimura, and Todd 1997) or doubly robust diff-in-diff (Sant'Anna and Zhao 2020), which incorporates both regression adjustment and inverse probability weighting.

This essay applies and expands on a weighting scheme suggested by Abadie (2005). Specifically, it uses the generalized maximum entropy logistic regression (GME Logit) as proposed by Golan Judge and Perloff (1996) to obtain the probability of selection into treatment and then uses the weighting scheme developed by Abadie to obtain the treatment effect. I show that the information theory (IT) based estimation of the selection equation outperforms against a classical logistic regression (Logit or ML Logit) for estimating selection probabilities in certain situations. This is because the IT-based estimator is known to outperform ML Logit in all finite samples. The more complex and ill-behaved the data, the larger the advantage of GME Logit.

Researchers are often interested in studying the treatment effect for a subpopulation (that is, for specific values of a covariate). Abadie (2005) also develops a semiparametric method to estimate this treatment effect for subpopulations, using the weights discussed above. Abadie's method replaces the true (usually high dimensional) conditional average treatment effect with a parametric approximation. If the parametric approximation is a linear model (often the case in practice), Abadie's method becomes weighted least squares (WLS) with the weights obtained from estimating the selection equation in stage 1. This essay expands on Abadie's method in this area as well, replacing the WLS-based estimator he suggests with a (weighted) generalized maximum entropy-based estimator. I also show that this method outperforms against the WLS-based estimator under certain circumstances.

The essay proceeds as follows: section II reviews the literature on diff-in-diff and conditional diff-in-diff; section III develops the theory; section IV compares the GME-based and classical methods in a series of simulations; section V applies the method to a well-known dataset and section VI concludes.

1 Review of Literature

The history of canonical diff-in-diff was briefly discussed earlier. I do not recap it here but discuss developments since then. The diff-in-diff literature developed in two distinct waves (Cunningham 2023). The first wave, which started in the 1980s (and discussed earlier), introduced the term and saw the publication of several influential papers. It peaked around 2007. The second wave began in 2011 and was characterized by "an exponential growth in popularity" (Cunningham 2023) and consistent use of terminology (such as parallel-trends and event study) in many papers.

This second wave continues today, with diff-in-diff becoming one of the most popular methods of causal identification.

The validity of what we now call the parallel trends assumption was under scrutiny very early in the history of the literature. For example, Ashenfelter (1978) discussed that participants in job training programs are often selected after they receive a temporary dip in their earnings or employment status. In such cases, wages could be expected to bounce back even in the absence of the training program. LaLonde (1986) showed that econometric methods such as diff-in-diff fail to match results obtained from a randomized control trial (RCT), and attributed this to mismatches in outcome trends between the treatment group and potential comparison groups (e.g. drawn from the current population survey, panel study of income dynamics or similar data sources for the general labor force) researchers might use. Heckman, et. al. (1998; 1997) proposed a method called regression adjustment that can recover the ATT when parallel trends are violated if the researcher correctly specifies the model for conditional outcome evolution. Recent research has progressed in three directions: (1) testing (conditional) parallel trends assumptions (e.g. (Roth 2022; Rambachan and Roth 2023)) (2) extending the method to staggered/multi-period treatment adoption (e.g. (Goodman-Bacon 2021; L. Sun and Abraham 2021; Callaway and Sant'Anna 2021)) or where the level/intensity of treatment can be varied continuously (e.g. (Callaway, Goodman-Bacon, and Sant'Anna 2024)) and (3) relaxing the parallel trends assumption (e.g. (Abadie 2005; Athey and Imbens 2006)). This essay is an expansion of methods that relax the parallel trends assumption.

2 Theory

I begin by defining the diff-in-diff methodology in the canonical setting. The setup is described in many textbooks such as (Cunningham 2021) or (Angrist and Pischke 2009). The researcher selects a causal quantity of interest (target parameter) for study. Adapting the potential outcomes framework of Rubin (1976), let $Y_{i,t}(0,0)$ denote unit $i's$ potential outcome at time t if it remains untreated in both periods, and $Y_{i,t}(0,1)$ if it is untreated in period $t = 0$ and treated in period $t = 1$. Since all units are untreated in period $t = 0$, we can simplify notation to: $Y_{i,t}(0) = Y_{i,t}(0,0)$ and $Y_{i,t}(1) = Y_{i,t}(0,1)$. Ideally, we would like to observe $Y_{i,t}(0)$ and $Y_{i,t}(1)$ for the same units. Then the treatment effect for unit i would be $Y_{i,t}(1) - Y_{i,t}(0)$. In practice, we can only observe $Y_{i,t}(1)$

for the treated units and $Y_{j,t}(0)$ for the untreated units. If D_i is an indicator variable that takes a value $D_i = 1$ if unit i is treated in period $t = 1$, and $D_i = 0$ otherwise, then the observed values are:

$$Y_{i,t} = (1 - D_i)Y_{i,t}(0) + D_i Y_{i,t}(1)$$

(1)

The potential outcomes framework defines a treatment effect $Y_{i,t}(1) - Y_{i,t}(0)$ for every unit i. In practice, researchers are often interested in finding the average treatment effect on the treated:

$$ATT = \mathbb{E}\left[Y_{i,t}(1) - Y_{i,t}(0)\middle|D_i = 1\right] = \mathbb{E}\left[Y_{i,t}(1)\middle|D_i = 1\right] - \mathbb{E}[Y_{i,t}(0)|D_i = 1]$$

(2)

Without additional assumptions, the ATT is not identified since we are unable to observe $\mathbb{E}\left[Y_{i,t}(0)\middle|D = 1\right]$ (the expected potential outcome if unit i had not received treatment, given that it actually did receive treatment). If we are willing to make additional assumptions, we could potentially identify the ATT. Different research designs make different assumptions for identifying the missing counterfactual. In the case of diff-in-diff, the crucial assumption (known as the parallel trends assumption) is:

$$\mathbb{E}\left[Y_{i,t=1}(0)\middle|D_i = 1\right] - \mathbb{E}\left[Y_{i,t=0}(0)\middle|D_i = 1\right] = \mathbb{E}\left[Y_{i,t=1}(0)\middle|D_i = 0\right] - \mathbb{E}\left[Y_{i,t=0}(0)\middle|D_i = 0\right]$$

(3)

In words, the expected change in outcome between pre-treatment and post-treatment periods ($\mathbb{E}\left[Y_{i,t=1}(0) - Y_{i,t=0}(1)\right]$) is the same for treated and comparison groups. Under these assumptions, the ATT is:

$$ATT = \mathbb{E}\left[Y_{i,t=1}\middle|D_i = 1\right] - \left(\mathbb{E}\left[Y_{i,t=0}\middle|D_i = 1\right] + \left(\mathbb{E}\left[Y_{i,t=1}\middle|D_i = 0\right] - \mathbb{E}\left[Y_{i,t=0}\middle|D_i = 0\right]\right)\right)$$

(4)

where the term in parenthesis on the right-hand side is our counterfactual $\mathbb{E}\left[Y_{i,t}(0)\middle|D_i = 1\right]$. The above can be rewritten as:

$$ATT = \left(\left[Y_{i,t=1}\middle|D_i = 1\right] - \mathbb{E}\left[Y_{i,t=0}\middle|D_i = 1\right]\right) - \left(\mathbb{E}\left[Y_{i,t=1}\middle|D_i = 0\right] - \mathbb{E}\left[Y_{i,t=0}\middle|D_i = 0\right]\right)$$

(5)

where the first term in the parenthesis is the change in the average outcome for treated units between period $t = 0$ and $t = 1$ and the second term is the change in the average outcome for the untreated units during the same period. Thus, the average treatment effect (on the treated) has a natural interpretation as the "additional" average change in the outcome of the treated units on top of the change experienced by the untreated units during the period of study.

It is well-known (for a textbook treatment, see (Cunningham 2021)) that estimating this model, involving four means is the same as a two-way fixed effects (TWFE) regression:

$$Y_{i,t} = \theta_t + \eta_i + \beta \cdot \mathbf{1}\{\mathrm{t} = 1\} \cdot \mathrm{D_i} + e_{i,t} \tag{6}$$

where θ_t and η_i are time and unit fixed effects respectively, $e_{i,t}$ are idiosyncratic shocks and β is the parameter of interest, $\mathbf{1}\{t = 1\}$ is an indicator variable if the observation is in the post-treatment period and D_i is an indicator variable if the unit received treatment. ($\mathbf{1}\{t = 1\} \cdot D_i = 1$ only for treated units in period $t = 1$ and 0 otherwise).

Often the parallel trends assumption is not plausible. The treatment group might differ from the comparison group in systematic ways that affect the outcome (for example, workers participating in a job training program might be less educated or younger than the comparison group and would have experienced a different wage trend than older workers regardless of whether they participated in training or not). We can relax the parallel trends assumption from above with a conditional parallel-trends assumption. Let X_i be a vector of observed determinants of changes in $Y_{i,t}$. We now assume that the groups would have followed parallel trends *conditional* on the covariates:

$$\mathbb{E}\left[Y_{i,t=1}(0) - Y_{i,t=0}(0) \middle| X_i, D_i = 1\right] = \mathbb{E}\left[Y_{i,t=1}(0) - Y_{i,t=0}(0) \middle| X_i, D_i = 0\right] \tag{7}$$

There are several ways to estimate this conditional parallel-trends model. Firstly, the covariates can be incorporated into an augmented TWFE regression (Baker et al. 2025) either as:

$$Y_{i,t} = \theta_t + \eta_i + \beta \cdot \mathbf{1}\{\mathrm{t} = 1\} \cdot \mathrm{D_i} + X_{i,t}\beta_{\mathrm{covs}} + e_{i,t} \tag{8}$$

or as

$$Y_{i,t} = \theta_t + \eta_i + \beta \cdot \mathbf{1}\{\mathrm{t} = 1\} \cdot \mathrm{D_i} + \left(\mathbf{1}\{t = 1\}X_{i,t=0}\right)\widetilde{\beta_{\mathrm{covs}}} + e_{i,t} \tag{9}$$

where $X_{i,t}$ is the observed value of the covariate at time t, $\mathbf{1}\{t = 1\}$ is an indicator variable that takes a value of 1 in the post-treatment period and 0 otherwise, and β_{covs} and $\widetilde{\beta_{covs}}$ are vectors of coefficients associated with the covariates. Note that since time-invariant variables drop out in the TWFE regression, only the effect of changes in the levels of a covariate $\Delta X_{i,t} = \left(X_{i,t=1} - X_{i,t=0}\right)$

for equation (8) or differential trends related to baseline levels of the variable $\mathbf{1}\{t = 1\}X_{i,t=0}$ for equation (9) can be estimated with this type of model. In addition, researchers need to be careful about what variables are controlled for. For example, if the treatment affects the outcome by changing the value of $X_{i,t}$ (e.g. job training improving some measured skill), then controlling for this in the TWFE regression will bias our estimate of ATT.

Abadie (2005) provides another way to estimate the average treatment effect. If we are willing to make the mild assumptions that (1) at least some portion of the population receives treatment and (2) at every stratum/value of the covariate, some portion of the population remains untreated, that is:

$$\Pr(D_i = 1) > 0 \ \textit{and} \ 0 < \Pr(D_i = 1|X_i) < 1$$

(10)

then Abadie (2005) shows that the ATT *conditional* on X_i is:

$$\mathbb{E}\left[Y_{i,t=1}(1) - E_{i,t}(0)\middle|X_i, D_i = 1\right] = \mathbb{E}\left[\rho_0 \cdot \left(Y(1) - Y(0)\right)\middle|X_i\right]$$

(11)

where:

$$\rho_0 = \frac{D_i - \Pr(D_i = 1|X_i)}{\Pr(D_i = 1|X_i)\,(1 - \Pr(D_i = 1|X_i)}$$

(12)

The (unconditional) ATT can be recovered as:

$$\mathbb{E}\left[Y_{i,t}(1) - Y_{i,t}(0)\middle|D_i = 1\right] = \int \mathbb{E}\left[\rho_0 \cdot \left(Y_{i,t}(1) - Y_{i,t}(0)\right)\middle|X_i\right] d\Pr(X_i|D_i = 1)$$

$$= \mathbb{E}\left[\frac{Y_{i,t}(1) - Y_{i,t}(0)}{\Pr(D_i = 1)} \cdot \frac{D_i - \Pr(D_i = 1|X_i)}{1 - \Pr(D_i = 1|X_i)}\right]$$

(13)

Abadie does not specify how $\Pr(D_i = 1|X_i)$ should be estimated. Commonly, researchers use ML Logit. However, other methods that estimate the selection probability could also be used. The GME approach for binary discrete choice (Golan, Judge, and Perloff 1996) is known to have several advantages over traditional ML Logit: it is efficient for small samples, avoids strong parametric assumptions, handles multicollinearity and is resilient to ill-conditioned data. The theory of GME Logit is recapped below (see Golan, Judge and Perloff 1996 for a more thorough development).

Assume that N units are observed where each unit is either selected into treatment or not. Let

$$p_i = \Pr(D_i = 1 | X_i, \gamma) = F(X_i'\gamma)$$

(14)

be the probability of observing unit $i \in [1,2, \dots N]$ in the treatment state, conditional on covariates X_i (of dimension $1 \times M$) and unknown parameters γ (of dimension $M \times 1$). If we have noisy data, we can write

$$y_i = F(\cdot) + e_i = p_i + e_i$$

(15)

where $y_i \in [0,1]$ is the observed outcome, p_i denotes the unknown and unobservable probability of selection and e_i is an error or noise component for each observation contained in $[-1, 1]$. We reparametrize the error component as:

$$e_i \stackrel{\text{def}}{=} \sum_h v_{ih} w_{ih}$$

(16)

where $\sum_h w_{ih} = 1$ are ~~an~~ H-dimensional vector of weights and v_i are the H-dimensional support space. (Golan, Judge, and Perloff 1996) suggest $v_i = \left[\frac{-1}{\sqrt{N}}, 0, \frac{1}{\sqrt{N}}\right]$ and thus $H = 3$.

We now maximize the entropy of the error augmented probability distribution:

$$\max_{p,w} \left(-\sum_i p_i \ln p_i - \sum_{ih} w_{ih} \ln w_{ih} \right)$$

(17)

subject to the M moment constraints imposed by the data, where the m-th constraint is:

$$\sum_i y_i X_{im} = \sum_i X_i p_i + \sum_{ih} X_i v_h w_{ih}$$

(18)

and the adding up constraint

$$\sum_h w_{ih} = 1$$

(19)

(I omit the constraint the $\sum p_{i,j}$=1 present in a multinomial version of this model, since that requirement is naturally satisfied in the case of binary choice). We can solve the above using the method of Lagrange multipliers. In matrix-form:

$$\mathcal{L} = -\boldsymbol{p}\ln p - \boldsymbol{w}'\ln \boldsymbol{w} + \boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{p} + \boldsymbol{X}'\boldsymbol{V}\boldsymbol{w} - \boldsymbol{X}'\boldsymbol{y}) + \boldsymbol{\mu}'(1 - I_T\boldsymbol{p}) + \boldsymbol{\rho}'(1 - \mathbf{1}'w)$$

(20)

where $\boldsymbol{\lambda}.\,\boldsymbol{\mu},\boldsymbol{\rho}$ are Lagrange multipliers and $\boldsymbol{\gamma} = -\boldsymbol{\lambda}$. The first order conditions are:

$$\frac{\partial \mathcal{L}}{\partial p_i} = -\ln p_i - 1 - \sum_m \lambda_m X_{im} - \mu_i = 0$$

$$\frac{\partial \mathcal{L}}{\partial w_{ih}} = -\ln w_{ih} - 1 - \sum_m \lambda_m X_{im} v_h - \rho_i = 0$$

$$\frac{\partial \mathcal{L}}{\partial w_{ih}} = \sum_i y_i X_{im} - \sum_i X_{im} p_i - \sum_k X_{im} v_h w_{ih} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \rho_i} = 1 - \sum_h w_{ih} = 0$$

(21)

From which we obtain the estimated probability distributions:

$$\widehat{p_i} = \frac{\exp\left(-\sum \widehat{\lambda_m} X_{im}\right)}{1 + \exp\left(-\sum_m \widehat{\lambda_m} X_{im}\right)}$$

(22)

and

$$w_{ih} = \frac{\exp\left(-\sum_m X_{im}\widehat{\lambda_m} v_h\right)}{\sum_h \exp\left(-\sum_m X_{im}\widehat{\lambda_m} v_h\right)}$$

(23)

Note that the above can be computed using a "dual unconstrained" method that is usually less computationally intensive but not discussed here for brevity (see (Golan 2017) for details).

3 Simulation

To evaluate the performance of alternative estimators of the ATT, I conduct Monte Carlo simulations based on a stylized data generating process (DGP) designed to reflect common features of observational studies, including selection into treatment on observables and covariate-dependent potential outcomes.

I begin by describing the baseline simulation. Since I am describing the data generating process for each unit i, I drop the i subscripts in this section. The simulated dataset contains $N = 500$ observations and the simulation is replicated across 5,000 Monte Carlo runs. Each observation is characterized by two baseline covariates, X_1 and X_2, independently drawn from a standard normal distribution (note that X_1 and X_2 represent two vectors of covariates, rather than unit or time subscripts):

$$X_1 \sim \mathcal{N}(0,1)$$
$$X_2 \sim \mathcal{N}(0,1)$$

(24)

The baseline outcome at $t = 0$, denoted Y_0, is modeled as a linear function of the covariates, with additive normal noise:

$$Y_0 = 5 + 0.5X_1 + 0.3X_2 + \epsilon_0,$$
$$\epsilon_0 \sim \mathcal{N}(0,1)$$

(25)

This structure ensures that baseline outcomes are systematically related to observed characteristics. In addition to level differences, the untreated trend also depends on the covariates:

$$\Delta_0 = 0.4X_1 - 0.1X_2$$

(26)

This trend is added to all individuals, regardless of treatment status, mimicking scenarios in which outcome trajectories differ across subgroups even in the absence of intervention. Treatment is assigned based on a logistic selection model that depends on the covariates:

$$\Pr(D = 1|X_1, X_2) = invlogit(0.5X_1 - 0.8X_2)$$

(27)

where $invlogit$ is $\Pr(D = 1|X1, X2) = \frac{\exp(0.5X_1 - 0.8X_2)}{\exp(0.5X_1 - 0.8X_2) + 1}$. (Recall $D = 1$) is an indicator variable on whether the unit received treatment). Each unit is assigned to treatment with a probability equal to this score. This introduces selection on observables, such that the probability of receiving treatment is systematically related to covariates that also influence potential outcomes. Equations

(25) and (26) are the crucial deviations from the unconditional parallel trends assumption that makes TWFE a biased estimator of the ATT. Since treatment into selection and the trend are both dependent on covariates X_1 and X_2, the treated and untreated groups no longer follow parallel trends.

The outcome in the post-treatment period ($t = 1$), denoted Y_1, is constructed as the untreated outcome plus the untreated trend, a constant treatment effect for the treated, and additional noise:

$$Y_1 = Y_0 + \Delta_0 + 2 \cdot D + \epsilon_1,$$
$$\epsilon_1 \sim \mathcal{N}(0,1)$$

(28)

Thus, the ATT is set to a constant value of 2 across all treated units, independent of covariates. However, since both the baseline outcome and the untreated trend depend on X_1 and X_2, failure to account for these covariates can lead to biased estimates of the treatment effect.

Under the baseline assumptions (large samples, well-behaved data), GME Logit and ML logit models of selection into treatment should behave similarly. To test whether GME Logit outperforms ML Logit under certain circumstances, I test several cases where GME is known to outperform. Firstly, GME is known to outperform ML logit for all finite samples. These differences are likely to be highlighted when sample sizes are very small. I modify the above (baseline) simulation, so that the dataset is of size $N = 25$ observations. Next, I test how the two estimators perform when the probability of receiving treatment is very small (this is done by shifting the selection equation to:

$$\Pr(D = 1|X_1, X_2) = invlogit(-3 + 0.4X_1 - 0.1X_2)$$

(29)

which causes only ~5% of observations to be selected into treatment). GME logit is also known to outperform ML logit when covariates are highly collinear. I simulate this by setting $X_2 = X_1 + u$ and $u \sim \mathcal{N}(0,0.01)$. Tables 1 and 2 summarize the simulation settings.

Table 16 Baseline simulation settings.

All simulations have 5,000 Monte Carlo runs.

Observations (N)	500
Covariate Distribution (X_1, X_2)	$X_1 \sim \mathcal{N}(0,1)$ $X_2 \sim \mathcal{N}(0,1)$
Heterogenous Trend (Δ_0)	$\Delta_0 = 0.4X_1 - 0.1X_2$
Selection equation	$\Pr(D = 1 \mid X_1, X_2) = invlogit(0.5X_1 - 0.8X_2)$
Errors	$\epsilon_0 \sim \mathcal{N}(0,1)$ $\epsilon_1 \mathcal{N}(0,1)$

Table 17 Modifications to the baseline case for specific scenarios.

All simulations have 5,000 Monte Carlo runs

Case: Small sample size	$N = 25$
Case: Rare selection	$\Pr(D = 1 \mid X_1, X_2) = invlogit(-3 + 0.5X_1 - 0.8X_2)$
Case: Highly collinear X_1, X_2	$X_2 = X_1 + u, u \sim \mathcal{N}(0,0.01)$

For the estimation of the GME Logit model, I use a support space of 3 with $v_{ih} = \left[\frac{-1}{\sqrt{n}}, 0, \frac{1}{\sqrt{n}}\right]$ as suggested by Golan, Judge and Perloff (1996). I do not use any prior distributions for GME Logit.

3.1 First Stage

I begin my analysis of the simulation results by first examining how successful ML Logit and GME Logit are at recovering the selection mechanism. I examine two aspects of the methods' performance: (1) how well each predicts treatment assignment (treated vs. untreated) and (2) how accurately each recovers the true model parameters. To assess (1), I report the number of misclassifications (following (Golan, Judge, and Perloff 1996)) as well as the Area Under the Receiver Operating Curve (AUC-ROC). The AUC-ROC is a widely used metric that summarizes a classifier's ability to rank positive cases above negative ones across all possible threshold values. A value of 0.5 corresponds to guessing at random, while a value of 1 indicates perfect classification. For reporting misses, I set the classifier threshold to 0.5 (that is, for each method, when the predicted probability is > 0.5, the observation is set to be in the treated group). For assessing (2), I report the Mean Squared Error (MSE) for each method as in Golan Judge and Perloff (1996).

Table 18 Model Performance for Stage 1 over 5,000 Monte Carlo simulation runs.

Each run has 500 observations, except for the small sample case, which has 25 observations. The worst, average and best misclassification rates are reported (misclassification is defin

		Misclassification Rate				
		Worst case	Average	Best Case	AUC	MSE
Baseline	GME	40.8%	33.1%	26.2%	0.73	0.02
	ML Logit	40.8%	33.1%	26.2%	0.73	0.02
Small Sample	GME	60.0%	29.1%	0%	0.76	0.24
	ML Logit	60.0%	29.1%	0%	0.76	625.48
Rare Selection	GME	11.2%	6.65%	3.0%	0.75	0.06
	ML Logit	11.2%	6.65%	3.0%	0.75	0.06
Highly Collinear	GME	50.6%	43.6%	35.6%	0.59	117.1
	ML Logit	50.6%	43.6%	35.6%	0.59	118.5

Table 3 shows the results for the first stage from the simulation. The results are similar to those obtained by Golan, Judge and Perloff (1996) in their exploration of this method. In the baseline case, GME Logit performs as well as ML Logit at classification, and has similar MSE. In addition to the metrics reported by them, I also include the AUC-ROC. These are also comparable between GME Logit and ML Logit.

Comparing GME Logit's performance with ML Logit for deviations from the baseline case, I note that both methods perform equally well at classification (whether looking at misclassification rates or AUC-ROC). Although the GME Logit has much lower MSE than ML Logit for small samples, these might not translate over to gains in the second stage of estimation. Only the estimate of the probability of selection into treatment (i.e. the propensity score) enters the second stage of estimation, and here both methods give similar results.

3.2 Second Stage

Recall that in my data generating process, the true treatment effect is "hard-coded" to a value of 2.0 (that is, a unit that receives treatment will have a 2.0 higher value for the outcome $Y_{i,t=1}$ if it receives treatment than it would otherwise). I now examine the ability of the second stage estimation (which simply involves replacing the weighted population expectations in equation 13 with their sample counterparts) to recover this treatment effect. In addition to the results obtained with weights derived from GME Logit and ML Logit first steps, I also show the results for the unweighted (or equal weighted) estimator. This estimator is biased and not consistent with the true ATT but is shown to give researchers an idea of what to expect when unconditional parallel trends assumption is violated in the underlying data, but the analysis does not account for the violation.

Table 19 Second Stage regression results.

See above for specification of each case ~~(leftmost column).~~ The true ATT is 2.0. Column headers indicate the model used for first stage propensity score estimation.

	MSE			Bias		
	GME	ML	Unweighted	GME	ML	Unweighted
Baseline	0.114	0.113	0.244	0.039	0.039	0.076
Small Sample	0.976	1.485	0.380	-0.739	-0.908	0.150
Rare Selection	0.073	0.074	0.149	-0.019	-0.027	0.280
Highly Collinear	0.032	0.032	0.018	0 .122	0.123	0 .038

Table 4 shows the results of the second stage. In general, ML Logit and GME Logit perform equally well (when measured by bias or MSE) in recovering the true treatment effect. This is even the case when the first stage GME estimator has lower MSE (e.g. for small sample sizes). In the second stage, the parameter estimates of the selection equation are not explicitly used. Only the propensity (likelihood of selection into treatment conditional on the covariates) enters the second stage weighting scheme. Golan, Judge and Perloff (1996) show that the GME estimator performs equally well with ML Logit when comparing misclassification rates but has lower variance and lower MSE. Thus, the results of the second stage are not unexpected. However, researchers might still prefer to use the GME Logit for the first stage if estimating the parameters of the selection mechanism is important, since GME Logit has lower variance than ML Logit.

3.3 Extension: Sub-population treatment effects

So far, our treatment effect has been constant (set to 2.0 in the previous examples) regardless of the value of the covariates. Thus, although the baseline values ($Y_0 = 5 + 0.5X_1 + 0.3X_2 + \epsilon_0$), trend ($\Delta_0 = 0.4X_1 - 0.1X_2$) and selection into treatment ($\Pr(D = 1|X_1, X_2) = invlogit(0.5X_1 - 0.8X_2)$ all depend on the covariates, the treatment effect for the entire population remains constant regardless of the values of X_1 and X_2. This might not be an accurate way to model how treatment works. Going back to the job training example, we might imagine that younger and more skilled workers could benefit more from training. If X_1 and X_2 are workers' ages and a proxy for skill, then the true treatment effect would be $\tau = g(X_1, X_2)$. Abadie's weighting method above would still recover the population-wide ATT, even if the treatment effect is dependent on covariates. However, researchers might be interested in studying the treatment effect for a substratum of the population (for example, for younger workers, for males vs. females, etc.). Abadie (2005) also suggests a way to estimate these subpopulation treatment effects, which is outlined below.

Consider X_k to be a deterministic function of the underlying covariates X (for example, X_k maybe a subset of X, or may contain indicator variables for X). Let $\mathcal{G} = \{g(X_k;\ \theta)|\theta \in \Theta \subset \mathbb{R}^k\}$ be a class of approximating functions, square integrable with respect to $\Pr(X_k|D = 1)$. Then a least squares approximation from $\mathcal{G}$ to $\mathbb{E}[Y_1(1) - Y_0(1)|X_k, D = 1]$ is given by $g(X_k; \theta_0)$, where:

$$\theta_0 = argmin_{\theta\in\Theta}\mathbb{E}[\{\mathbb{E}[Y_1(1) - Y_0(1)|X_k, D = 1] - g(X_k; \theta)\}^2|D = 1]$$

(30)

For example, if $\mathcal{G} = \{X_k'\theta|\theta \in \Theta \subset \mathbb{R}^k\}$ then θ_0 defines a linear LS approximation to $\mathbb{E}[Y_1(1) - Y_0(1)|X_k, D = 1]$. Abadie shows that the θ_0 can be recovered as:

$$\theta_0 = argmin_{\theta\in\Theta}\mathbb{E}\left[\Pr(D = 1|X) \cdot \left\{\rho_0 \cdot \big(Y(1) - Y(0)\big) - g(X_k; \theta)\right\}^2\right]$$

(31)

In the case when $\mathcal{G} = \{X_k'\theta|\theta \in \Theta \subset \mathbb{R}^k\}$, this simply becomes weighted LS (with weights obtained from the first stage).

It is known (see (Golan, Judge, and Miller 1996; Golan 2007) that the generalized maximum entropy linear model (GME Linear) is superior to LS for finite samples and ill-behaved data. Thus, it is natural to extend the weighted LS (WLS or weighted LS) proposed of Abadie to a weighted GME approach, to see how it performs. To test this, I replace the constant treatment effect $\tau = 2.0$ with

$$\tau = 2.0 + 0.5X_1 - 0.3X_2$$

(32)

Both weighted LS regression and weighted GME regression should be consistent estimators of the conditional average treatment effect (32). For the weights, I use the predicted selection probabilities obtained by performing a GME Logit in the first stage $w_i = \Pr(D_i = 1|X_i)$. Since the GME Logit function in Stata does not accept weights, I obtain GME results by multiplying each observation by $\sqrt{\omega_i}$.

For the GME Linear estimation, we also need to specify a support space for the parameters and a support space for the errors. Since the model is now in a "weighted space" (each observation weighted by $\sqrt{\omega_i}$), the empirical $\pm 3\sigma$ rule is no longer appropriate. I also use a wide support space for the parameters, to capture extreme parameter values. Equation (33) specifies the support spaces:

$$w_{ih} = [-100, -91.67, \dots 0, \dots, 91.67, 100]; \{h: [1,25]\}$$

$$Z = \begin{bmatrix} -100 & 0 & 100 \\ -100 & 0 & 100 \\ -100 & 0 & 100 \end{bmatrix}$$

(33)

The error support space for each observation w_i is an equally spaced vector of size 25, ranging from -100 to 100. The parameter support space is a 3×3 matrix with each parameter having the support space $[-100, 0, 100]$.

Table 5 shows weighted LS and OLS results in head-to-head comparison. In general, both methods behave similarly and are successful at obtaining the conditional average treatment effect. However, GME Logit has significantly lower variance in the case of highly collinear variables. The weighted LS regression has over 300 times the MSE as the GME estimator. Thus, researchers should prefer the GME based estimator since it performs as well as WLS in all circumstances and strictly better when covariates are highly collinear.

Table 20 Estimate of the conditional average treatment effect.

The true conditional treatment effect is $\tau(X_1, X_2) = 2.0 + 0.5X_1 - 0.3X_2$

	Bias						MSE					
	β_0		β_1		β_2		β_0		β_1		β_2	
	WLS	GME	WLS	GME	WLS	GME	WLS	GME	WLS	GME	WLS	GME
Baseline	0.001	-0.003	0.009	-0.001	-0.008	0.003	0.020	0.023	0.047	0.080	0.049	0.111
Small Sample	-0.005	-0.093	0.107	0.074	-0.133	-0.122	0.435	0.389	0.714	0.628	0.657	0.574
Rare Selection	-0.180	-0.195	0.003	0.009	-0.116	-0.119	0.176	0.176	0.155	0.151	0.146	0.143
Collinear	-0.001	-0.005	0.074	-0.634	-0.074	0.633	0.017	0.017	194.47	0.505	194.42	0.504

In general, researchers should prefer the info-metrics based methods for estimating selection probabilities (stage 1) and for estimating the conditional treatment effects. The info-metrics method outperforms ML Logit for small sample sizes and the info-metrics based conditional average treatment effect outperforms weighted LS when covariates are highly collinear.

4 Application to Real World Data

I now apply this method to a real-world dataset. I use the data provided by LaLonde (1986) for his evaluation of the impact of the National Supported Work (NSW) demonstration. The data is readily available in many software packages as well as through the NBER website. As described by LaLonde, the NSW program randomly assigned individuals to a treatment or control group. Those in the treatment group received "supported work" where participants were employed at construction, service or similar industries in a supportive but performance-oriented work environment. Those in the control group were left to fend for themselves. LaLonde examines the data for male and female participants separately. For brevity, I focus on the analysis for male participants only. I also only use the info-metrics based estimators (due to the reasons discussed earlier). For the male group, the outcome of interest is the 1978 real earnings due to participation in the program (or earnings growth from baseline 1975 wages in a difference-in-differences context). Since we have data from a control group, the difference in mean 1978 earnings between the experimental and control group is an unbiased estimator of the treatment effect. LaLonde asks, what if researchers used an alternative control group to analyze the data instead of the actual control group. He constructs several sets of control groups: one based on the Current Population Survey (CPS-1), one based on the Panel Study of Income Dynamics (PSID-1), and several others that are obtained by sub-setting CPS-1 or PSID-1 samples to match pre-treatment earnings and demographic characteristics of program participants. LaLonde's results with respect to this dataset (as well as using from the actual control group, which serves as the "gold standard" estimate of the treatment effect) is reproduced below.

Table 21 Replication of LaLonde (1986) results

Treatment effects for male NWS participants as reported in LaLonde (1986) using the actual control group, and CPS-1 group drawn from the current population survey with similar characteristics as the treatment group. Standard Errors in parenthesis. (Source: LaLonde (11986) and author's calculations)

Comparison Group	Treatment Less Comparison Group Earnings		Diff-in-Diff: Difference in Earnings Growth 1975 to 1978 Less Comparison Group	
	Unadjusted	Adjusted[9]	Without Age	With Age
Actual Controls	\$886 (476)	\$798 (472)	\$847 (560)	\$856 (558)
CPS-1	-\$8,870 (562)	-\$4,416 (577)	\$1,714 (452)	\$195 (441)

The gold standard point estimate of the ATT (as shown in row 1 of Table 6) is \$886 (or \$798 if we adjust for difference in age, schooling, etc. between treatment and comparison groups). LaLonde also performs a diff-in-diff analyses (this is strictly not needed when we are using actual RCT control group as our comparison group, since the simple difference in means between the control and treatment groups in the treatment period is an unbiased estimator of the ATT. However, a diff-in-diff is also a consistent and unbiased estimator of the ATT). LaLonde performs both a simple diff-in-diff and one that adjusts for the difference in age between the treatment and control group (other characteristics such as education and ethnicity remain constant for each individual across periods, so are not used in the diff-in-diff). Point estimates from each method range from \$798 (for the demographics adjusted difference in mean earnings) to \$886 (for

[9] The exogenous variables used in the regression adjusted equations are age, age squared, years of schooling, high school dropout status and race

unadjusted differences between the treatment and control groups), with the diff-in-diff estimators falling in between. At the $\alpha = 0.05$ level of significance, we cannot reject the null hypothesis that the treatment had no effect on earnings. This is in line with other analysis of the NWS program which finds no effect on earnings for male participants (e.g. (A. Smith and E. Todd 2005; Dehejia and and Wahba 1999)).

I now focus on row 2 of Table 6, which answers the question ‘if researchers did not have a RCT control group but used a comparison group such as the current population survey, how successful would they be at matching the “gold standard” ATT estimate’. The CPS-1 group consists basically of all male headed-households in the CPS who were less than age 55 in 1975 with matched social security numbers (Westat’s matching criteria). The naïve difference in 1978 earnings between the treatment group and CPS comparison group (column 1) recovers a large negative impact of participating in the program. Using the experimental control group as our reference, we know participation in the program is associated with a ~$886 increase in earnings. A naïve comparison between the CPS-1 sample’s wages and the treatment group’s wages is not successful at recovering the ATT. The large negative impact obtained when using the CPS sample as the comparison group is likely due to differences in demographic characteristics between the treatment group and the CPS sample. Column 2 attempts to adjust for some demographic characteristics such as educational attainment, race and age through a regression. However, the estimate of the program’s effect continues to be negative, likely due to uncontrolled for (and unobservable) differences between the CPS-1 group and NSW treatment group. Columns 3 and 4 show the results of a diff-in-diff type estimation (using 1975 earnings as a baseline). Results in column 4 control~~s~~ for age (and age squared) in the ~~DiD~~ estimate. The CPS group shows wildly different impact of program participation when age is controlled for, and none are close to our point estimate of the ATT.

I now apply Abadie’s weighting method, using the GME based first stage described earlier. For estimating the selection equation, I use the age, age squared, years of education, an indicator for whether the individual has a high school degree, an indicator for the individual’s race or ethnicity (white is baseline), and an indicator for whether the individual is married. I obtain standard errors through the bootstrap. I use the same error support weight vector as specified earlier: $\left[\frac{-1}{\sqrt{N}}, 0, \frac{+1}{\sqrt{N}}\right]$. Table 7 shows the results. The estimate of the program’s effect is now much closer to one obtained from using the experimental control group. The standard error of the estimate is also similar. If researchers used the CPS sample but weight the observations so that treatment and control groups look similar, they would obtain estimates and reach similar conclusions as they would using an RCT.

Table 22 Treatment effects using Abadie's weighting scheme (with GME first step).

The best point estimate of the treatment effect (using the control group from the RCT as the comparison group) is $847. Reweighting observations to account for selection into treatm

Comparison Group	Difference in Earnings
Controls	$847 (560)
CPS-1 with Abadie's method (GME first stage)	$575 (562)

5 VI. Conclusion

This paper advances the conditional difference-in-differences (DiD) framework by integrating an information-theoretic approach, specifically generalized maximum entropy (GME) methods, to enhance the estimation of treatment effects under violations of the parallel trends assumption. Through Monte Carlo simulations, I demonstrate that GME Logit matches or outperforms maximum likelihood logistic regression (ML Logit) in estimating selection probabilities, particularly in small samples and rare selection scenarios, due to its lower mean squared error and robustness to ill-conditioned data. While both methods yield comparable ATT estimates in the second stage, GME Logit's advantages in parameter estimation make it a valuable tool when modeling selection mechanisms is a priority. Furthermore, extending Abadie's (2005) weighted least squares approach to a weighted GME Linear estimator for subpopulation treatment effects shows promise in handling heterogeneous treatment effects, especially in finite samples or with complex data structures.

Applying the method to LaLonde's (1986) National Supported Work dataset, I find that the GME-based DiD estimator, using Abadie's weighting scheme, produces an ATT estimate ($575) closer to the randomized controlled trial benchmark ($886) than traditional ~~DiD~~ methods when using a non-experimental comparison group like CPS-1. This underscores the method's ability to mitigate biases from covariate imbalances, offering a robust alternative when conditional parallel trends hold.

These findings have implications for empirical research, particularly in observational studies where data limitations challenge standard assumptions. The GME-based approach provides a flexible and efficient framework for causal inference, accommodating small samples, rare events, and collinear covariates. Future research could explore its performance in multi-period or staggered treatment settings, integrate doubly robust methods, or develop computational optimizations for large-scale applications. By leveraging information-theoretic principles, this work contributes to the ongoing evolution of causal inference methods.

Appendix A Contract Characteristics

This appendix lists the detailed characteristics of the “standard” (most liquid) interest-rate swaps contract for the currencies studied in this essay.

Table 23 Contract characteristics for common contracts (Source: Bloomberg L.P.)

Currency	USD	CAD	GBP	CHF
Settlement	T+2	T+0	T+0	T+2
Fixed Leg				
Day Count Convention	30I/360	ACT/365.FIXED	ACT/365.FIXED	30E/360
Payment Frequency	Semiannual	Semiannual	Semiannual	Annual
Business Day Adjustment Convention	Modified Following	Modified Following	Modified Following	Modified Following
Adjustment Type	Accrual and Payment Dates	Accrual and Payment Dates	Accrual and Payment Dates	Accrual and Payment Dates
Roll Convention	Backward	Backward	Backward (EOM)	Backward
Accrual Calculation Calendar	US Federal Reserve and Bank of England	Canada	Bank of England	Switzerland
Pay Delay	0 days	0 days	0 days	0 days
Floating Leg				
Day Count Convention	Actual/360	ACT/365.FIXED	ACT/365.FIXED	Actual/360
Payment Frequency	Quarterly	Semiannual	Semiannual	Semiannual
Reference Index	USD LIBOR 3M	CDOR 3M	GBP LIBOR 6M	CHF LIBOR 6M
Reset Frequency	Quarterly	Quarterly	Semiannual	Semiannual
Business Day Adjustment	Modified Following Business Day	Modified Following	Modified Following	Modified Following

Adjustment Type	Accrual and Payment Dates	Accrual and Payment Dates	Accrual and Payment Dates	Accrual and Payment Dates
Roll Convention	Backward	Backward	Backward (EOM)	Backward
Calculation Calendar	US Federal Reserve, England	Canada	England	Switzerland
Fixing Calendar	England	Canada	England	England
Fixing Lag	2 business days	0 days	0 days	2 business days
Pay Delay	0 days	0 days	0 days	0 days
Reset Position	Advance	Advance	Advance	Advance

Definitions

Settlement

Settlement refers to the number of business days after the trade date when the swap contract is finalized, and payments are made (Hull 2022). The most common conventions are T+0, T+2, and T+3, where "T" represents the trade date, and the number indicates how many business days after the trade date settlement occurs. For example, in a T+2 settlement, the settlement occurs two business days after the contract is executed.

Fixed Leg

The fixed leg of an interest rate swap refers to the portion of the swap where the payer makes periodic payments at a fixed interest rate, which is predetermined and remains constant throughout the life of the swap (Hull 2022). The characteristics below describe various conventions associated with this leg.

- Day Count Convention (Ametrano and Ballabio 2003b): This convention determines how interest accrues over time, using fractions of a year based on the number of days between two dates . Common conventions include:
    - 30I/360: Assumes each month has 30 days and a year has 360 days. It simplifies calculations but may deviate slightly from actual time. If the start date of the day count is on 31st, it is treated as if it is the 30th. If the end date is on the 31st, then it can either be treated as the 30th or the 31st depending on the start date. It also

includes special rules for when either the start or end date of the day count convention is Feb 28/29.

  - 30E/360: Like the 30I/360 convention, but if either the start or end dates are on the 31st, they are treated as if they are on the 30th. February 28/29 is always treated as 30th.
  - ACT/365.FIXED: Uses the actual number of days in a period, dividing by a fixed 365-day year for calculating partial year interest rate accrual.
  - Actual/360: Uses the actual number of days between dates, but assumes the year has 360 days when calculating partial year interest payments.

- Payment Frequency (Ametrano and Ballabio 2003b): This defines how often payments are made on the fixed leg. For instance, "semiannual" means payments are made twice a year, while "annual" means once a year.
- Business Day Adjustment Convention (Ametrano and Ballabio 2003b): When a payment date falls on a non-business day, this convention dictates how the date is adjusted. All the contracts in this study use a "modified following" convention. A modified following convention means payments are pushed to the next business day unless that day falls in the next month, in which case payments are moved backward to the preceding business day.
- Adjustment Type (Ametrano and Ballabio 2003b): Adjustment type refers to which dates are adjusted when a business day adjustment is necessary. For example, in "Accrual and Payment Dates" adjustment, both the accrual period and the payment date will be adjusted if necessary.
- Roll Convention (Ametrano and Ballabio 2003b): The roll convention specifies how payment dates are set relative to a reference date, typically whether payments move forward or backward when adjusting for business days. A "Backward" roll moves the date to the nearest preceding business day, while "Backward (EOM)" additionally ensures payments align with end-of-month periods.
- Accrual Calculation Calendar (Ametrano and Ballabio 2003b): This calendar determines which set of business days are considered in calculating the accrual of interest payments. For example, the "US Federal Reserve" calendar includes only U.S. federal holidays, while the "England" calendar takes U.K. public holidays into account.

- Pay Delay (Bloomberg L.P.): Pay delay refers to the number of days between the payment due date and the actual date the payment is made. For instance, "0 days" means payments are made on the due date.

Floating Leg

The floating leg of the swap is where payments are made based on a variable interest rate, which changes over time based on a reference index (Hull 2022). The conventions below describe how these payments are structured.

- Reference Index (Hull 2022; Ametrano and Ballabio 2003b): The reference index is the benchmark interest rate that dictates the floating payments. Common indices include:
    - USD LIBOR 3M: U.S. Dollar London Interbank Offered Rate for a 3-month period.
    - CDOR 3M: Canadian Dollar Offered Rate for 3 months.
    - GBP LIBOR 6M: British Pound LIBOR for 6 months.
    - CHF LIBOR 6M: Swiss Franc LIBOR for 6 months.
- Reset Frequency (Ametrano and Ballabio 2003b): This determines how often the floating rate is recalculated or "reset." For example, a quarterly reset means the floating rate is updated every three months.
- Fixing Calendar (Hull 2022; Ametrano and Ballabio 2003b): This refers to the calendar used to determine when the floating rate is fixed or set. For example, the "England" fixing calendar means rates are set according to U.K. business days.
- Fixing Lag: Fixing lag defines how many days in advance the floating rate is determined before the payment period begins. For instance, a "2 business days" fixing lag means the floating rate is set two days before the payment is due.

Reset Position: "Advance" reset position means the floating rate is set at the beginning of the interest period and applied throughout the period.

Appendix B Robustness Tests

This appendix provides several robustness tests to my analysis. Firstly, in the main body of the essay, in the analysis of the clearing mandate on swap pricing, I filtered out observations that were +/- 50 bps from the Bloomberg terminal calculated fair rate. I now present an alternate version of tables Table 8 and Table 9, now including these outliers. Overall, there were 1,101 such outliers (representing about 4% of the overall dataset). Results continue to be very similar to the results found in the main essay. In this broader dataset, clearing causes a 12-bps rise in swaps prices for USD contracts in the overall dataset. Most control variables also show similar results to what is found in the main essay. The notable exceptions are tenor (where a one-year increase in the tenor is now associated with a 0.03 bps increase in the premium instead of a 4-bps decrease) and Friday trading (which is now associated with a 0.95 bps increase in the premium rather than a -0.78-bps decrease). The group difference (that is the difference in baseline premium for USD over CAD contracts) also becomes not statistically significant. Note that the effect of these control variables is small (less than 1 bps).

Table 24 Diff-in-diff results after including all pricing outliers

Diff-in-diff Regression Results

	Dependent variable: Premium	
	Basic Model	Advanced Model
	(1)	(2)
Group	6.4783***	7.1186***
	(1.2754)	(1.2786)
Period	-16.8364***	-16.6251***
	(1.6893)	(1.6921)
Tenor		-0.0552**
		(0.0226)
Log Notional		0.5606***

		(0.1730)
Capped		-0.3432
		(0.4880)
SEF		-4.2437
		(6.5996)
Morning Session		-2.6170***
		(0.4843)
Afternoon Session		-2.4587***
		(0.4762)
Off Hours		-3.5615***
		(0.5589)
Monday		2.8288***
		(0.5910)
Tuesday		1.6594***
		(0.5454)
Thursday		1.1212**
		(0.5274)
Friday		-0.7841
		(0.5581)
Group * Period	12.8246***	12.1513***
	(1.7395)	(1.7539)
Constant	-3.4148***	-11.6213***
	(1.2252)	(3.1911)
Observations	28,311	28,311
R^2	0.0139	0.0182
Adjusted R^2	0.0138	0.0177

Residual Std. Error	30.2838 (df = 28307)	30.2239 (df = 28296)
F Statistic	133.0846*** (df = 3; 28307)	37.4453*** (df = 14; 28296)
Note:		*p**p***p<0.01

Table 25 Difference-in-differnces results after including pricing outliers. By phase analysis.

By Phase Results: Advanced Model

	Dependent variable: Premium		
	Phase 1	Phase 2	Phase 3
	(1)	(2)	(3)
Group	7.200***	6.368***	12.806***
	(2.241)	(2.151)	(2.331)
Period	-2.813	-16.738***	-5.274**
	(3.751)	(2.926)	(2.617)
Tenor	-0.305***	0.021	0.039
	(0.055)	(0.031)	(0.035)
Notional	-2.382***	0.653**	2.108***
	(0.391)	(0.257)	(0.257)
Capped	1.164	0.031	-1.834**
	(1.141)	(0.689)	(0.729)
Morning Session	-5.318***	0.581	-2.231***
	(1.121)	(0.702)	(0.716)
Afternoon Session	-3.678***	-1.300*	-0.882
	(1.110)	(0.673)	(0.718)
Off Hours	-6.020***	0.474	-6.191***

	(1.311)	(0.803)	(0.826)
Monday	-0.927	9.518^{***}	-4.437^{***}
	(1.361)	(0.883)	(0.866)
Tuesday	-2.788^{**}	8.744^{***}	-3.307^{***}
	(1.315)	(0.787)	(0.794)
Thursday	-2.880^{**}	9.627^{***}	-5.501^{***}
	(1.292)	(0.736)	(0.790)
Friday	-2.307^{*}	4.595^{***}	-4.177^{***}
	(1.373)	(0.768)	(0.848)
Group * Period	-3.568	11.873^{***}	11.274^{***}
	(3.848)	(3.007)	(2.755)
Constant	50.354^{***}	-21.877^{***}	-50.813^{***}
	(6.877)	(4.958)	(4.895)
Observations	7,819	11,233	9,259
R^2	0.025	0.036	0.073
Adjusted R^2	0.023	0.035	0.071
Residual Std. Error	37.150 (df = 7805)	26.916 (df = 11219)	25.619 (df = 9245)
F Statistic	15.137^{***} (df = 13; 7805)	32.443^{***} (df = 13; 11219)	55.763^{***} (df = 13; 9245)
Note:			*p**p***p<0.01

Also, as is sometimes done in literature, I drop observations on the first trading day of each phase that the clearing mandate went into force (to mitigate the effects of program implementation effects). Results continue to show similar patterns as found in the main body of the essay. The mandate causes a 12-bps rise in premiums for US-contracts after implementation. Control variables show similar signs and magnitude as is discussed in the main body of the paper.

Table 26 Diff-in-diff results excluding the first day the treatment went into effect

Diff-in-diff Regression Results

	Dependent variable: Premium	
	Basic Model	Advanced Model
	(1)	(2)
Group	-0.8889*	-0.7596
	(0.4942)	(0.4928)
Period	-11.4254***	-11.1961***
	(0.7048)	(0.7007)
Tenor		0.0305***
		(0.0089)
Log Notional		0.6367***
		(0.0693)
Capped		-0.7166***
		(0.1920)
SEF		0.4531
		(2.5989)
Morning Session		-0.9881***
		(0.1916)
Afternoon Session		-1.1930***
		(0.1877)
Off Hours		-1.2874***

		(0.2193)
Monday		2.6462***
		(0.2662)
Tuesday		2.3934***
		(0.2081)
Thursday		2.7678***
		(0.2016)
Friday		0.9347***
		(0.2137)
Group * Period	12.0888***	11.5072***
	(0.7234)	(0.7226)
Constant	-0.2415	-11.9859***
	(0.4742)	(1.2767)
Observations	25,613	25,613
R^2	0.0175	0.0335
Adjusted R^2	0.0174	0.0330
Residual Std. Error	11.4106 (df = 25609)	11.3196 (df = 25598)
F Statistic	152.0776*** (df = 3; 25609)	63.4120*** (df = 14; 25598)
Note:		*p**p***p<0.01

Finally, given the low R^2 values of some of the regression results, I try alternative regression specifications. Table 27 shows the results of a model with second-order terms for the continuous control variables tenor and notional. Although these higher order terms are statistically significant, the model still suffers from the same low R^2 as the model used in the main body of the essay. The overall conclusion remains the same (clearing causes a 13-bps increase in premium for USD contracts). Control variables also show similar signs and magnitudes as in the main body of the essay.

Table 27 Diff-in-diff with alternative model specifications

Diff-in-diff Regression Results

	Dependent variable: Premium
	Advanced Model
Group	-0.6567
	(0.4897)
Period	-13.3189***
	(0.6611)
$Tenure^2$	0.0007***
	(0.0003)
$(Ln\ notional)^2$	0.0235***
	(0.0021)
Capped	-1.0574***
	(0.1883)
SEF	0.7100
	(2.5202)
Morning	-1.0268***
	(0.1843)
Afternoon	-1.2300***
	(0.1814)
Off Hours	-1.3042***
	(0.2125)
Monday	1.5552***
	(0.2245)
Tuesday	2.3847***
	(0.2071)

Thursday	2.7650^{***}
	(0.2006)
Friday	0.9402^{***}
	(0.2125)
Group * Period	13.4560^{***}
	(0.6840)
Constant	-7.6601^{***}
	(0.7760)
Observations	27,210
R^2	0.0441
Adjusted R^2	0.0436
Residual Std. Error	11.2629 (df = 27195)
F Statistic	89.5578^{***} (df = 14; 27195)
Note:	*p**p***p<0.01

REFERENCES

A. Smith, Jeffrey, and Petra E. Todd. 2005. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics*, Experimental and non-experimental evaluation of economic policy and models, 125 (1): 305–53. https://doi.org/10.1016/j.jeconom.2004.04.011.

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *The Review of Economic Studies* 72 (1): 1–19.

Abadie, Alberto, and Guido W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74 (1): 235–67. https://doi.org/10.1111/j.1468-0262.2006.00655.x.

Allen, Franklin, and Douglas Gale. 2000. "Financial Contagion." *Journal of Political Economy* 108 (1): 1–33. https://doi.org/10.1086/262109.

Allen, Franklin, and Douglas M. Gale. 2007. "An Introduction to Financial Crises." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.1008311.

Ametrano, Ferdinando, and Luigi Ballabio. 2003a. "The QuatLib Library: An Open Source Framework for Quantitative Finance." http://www.quantlib.org.

———. 2003b. "User's Manual for QuantLib: An Open Source Framework for Quantitative Finance." User's Manual 0.3.1 beta. http://www.quantlib.org: TheOrganisation.

Angrist, Joshua D. 1990. "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records." *The American Economic Review* 80 (3): 313–36.

———. 1991. "The Draft Lottery and Voluntary Enlistment in the Vietnam Era." *Journal of the American Statistical Association* 86 (415): 584–95. https://doi.org/10.1080/01621459.1991.10475083.

Angrist, Joshua D., and Guido W. and Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430): 431–42. https://doi.org/10.1080/01621459.1995.10476535.

Angrist, Joshua D., and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?*." *The Quarterly Journal of Economics* 106 (4): 979–1014. https://doi.org/10.2307/2937954.

———. 1999. "Chapter 23 - Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, 3:1277–1366. Elsevier. https://doi.org/10.1016/S1573-4463(99)03004-7.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

———. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." *Journal of Economic Perspectives* 24 (2): 3–30. https://doi.org/10.1257/jep.24.2.3.

Ashenfelter, Orley. 1978. "Estimating the Effect of Training Programs on Earnings." *The Review of Economics and Statistics* 60 (1): 47–57. https://doi.org/10.2307/1924332.

Ashenfelter, Orley, and David Card. 1984. "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w1489.

Athey, Susan, and Guido W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74 (2): 431–97. https://doi.org/10.1111/j.1468-0262.2006.00668.x.

Azad, A. S. M. Sohel, Victor Fang, and Chi-Hsiou Hung. 2012. "Linking the Interest Rate Swap Markets to the Macroeconomic Risk: The UK and Us Evidence." *International Review of Financial Analysis* 22 (April):38–47. https://doi.org/10.1016/j.irfa.2012.03.001.

Baker, Andrew, Brantly Callaway, Scott Cunningham, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2025. "Difference-in-Differences Designs: A Practitioner's Guide." arXiv. https://doi.org/10.48550/arXiv.2503.13323.

Bebchuk, Lucian A., and Itay Goldstein. 2011. "Self-Fulfilling Credit Market Freezes." *The Review of Financial Studies* 24 (11): 3519–55. https://doi.org/10.1093/rfs/hhr086.

Benos, Evangelos, Wenqian Huang, Albert J. Menkveld, and Michalis Vasios. 2019. "The Cost of Clearing Fragmentation." SSRN Scholarly Paper ID 3502465. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3502465.

Benos, Evangelos, Richard Payne, and Michalis Vasios. 2020. "Centralized Trading, Transparency, and Interest Rate Swap Market Liquidity: Evidence from the Implementation of the Dodd–Frank Act." *Journal of Financial and Quantitative Analysis* 55 (1): 159–92. https://doi.org/10.1017/S0022109018001527.

Bernstein, Asaf, Eric Hughson, and Marc Weidenmier. 2019. "Counterparty Risk and the Establishment of the New York Stock Exchange Clearinghouse." *Journal of Political Economy* 127 (2): 689–729. https://doi.org/10.1086/701033.

Biais, Bruno. 1993. "Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets." *The Journal of Finance* 48 (1): 157–85. https://doi.org/10.1111/j.1540-6261.1993.tb04705.x.

Bicksler, James, and Andrew H. Chen. 1986. "An Economic Analysis of Interest Rate Swaps." *The Journal of Finance* 41 (3): 645–55. https://doi.org/10.2307/2328495.

Bloomberg L.P. 2013. "Yield Curve Data for U.S. Swap Rates and Canadian Dollar Rates." Bloomberg Terminal.

———. 2024. "Building the Bloomberg Interest Rate Curve – Definitions and Methodology." Bloomberg L.P.

Bolandnazar, Mohammadreza. 2020. "The Market Structure of Dealer-to-Client Interest Rate Swaps." SSRN Scholarly Paper 3752367. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3752367.

Boudiaf, Ismael Alexander, Immo Frieden, and Martin Scheicher. 2024. "The Market Liquidity of Interest Rate Swaps." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4745740.

Brunnermeier, Markus K. 2009. "Deciphering the Liquidity and Credit Crunch 2007-2008." *Journal of Economic Perspectives* 23 (1): 77–100. https://doi.org/10.1257/jep.23.1.77.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "Difference-in-Differences with a Continuous Treatment." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w32117.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001.

Capponi, Agostino, and Martin Larsson. 2015. “Price Contagion through Balance Sheet Linkages.” *The Review of Asset Pricing Studies* 5 (2): 227–53. https://doi.org/10.1093/rapstu/rav006.

Cifuentes, Rodrigo, Gianluigi Ferrucci, and Hyun Song Shin. 2005. “Liquidity Risk and Contagion.” *Journal of the European Economic Association* 3 (2–3): 556–66. https://doi.org/10.1162/jeea.2005.3.2-3.556.

Commodity Futures Trading Commission. n.d. “Dodd-Frank Act.” CFTC.Gov. Accessed June 30, 2025. https://www.cftc.gov/LawRegulation/DoddFrankAct/index.htm.

Cross, Frank. 1973. “The Behavior of Stock Prices on Fridays and Mondays.” *Financial Analysts Journal* 29 (6): 67–69. https://doi.org/10.2469/faj.v29.n6.67.

Cunningham, Scott. 2021. *Causal Inference: The Mixtape*. New Haven London: Yale University Press.

———. 2023. “The First and Second Waves of Difference-in-Differences.” Substack newsletter. *Scott’s Mixtape Substack* (blog). November 9, 2023. https://causalinf.substack.com/p/the-first-and-second-waves-of-difference.

Darbyshire, J. H. M. 2022. *Pricing and Trading Interest Rate Derivatives: A Practical Guide to Swaps*. Aitch & Dee Limited.

Dehejia, Rajeev H., and Sadek and Wahba. 1999. “Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs.” *Journal of the American Statistical Association* 94 (448): 1053–62. https://doi.org/10.1080/01621459.1999.10473858.

Diamond, Douglas W., and Philip H. Dybvig. 1983. “Bank Runs, Deposit Insurance, and Liquidity.” *Journal of Political Economy* 91 (3): 401–19. https://doi.org/10.1086/261155.

Diamond, Douglas W., and Raghuram G. Rajan. 2011. “Fear of Fire Sales, Illiquidity Seeking, and Credit Freezes *.” *The Quarterly Journal of Economics* 126 (2): 557–91. https://doi.org/10.1093/qje/qjr012.

Dodd-Frank Wall Street Reform and Consumer Protection Act. 2010. *United States Code*.

Duffie, Darrell, Martin Scheicher, and Guillaume Vuillemey. 2015. “Central Clearing and Collateral Demand.” *Journal of Financial Economics* 116 (2): 237–56. https://doi.org/10.1016/j.jfineco.2014.12.006.

Duffie, Darrell, and Haoxiang Zhu. 2011. “Does a Central Clearing Counterparty Reduce Counterparty Risk?” *Review of Asset Pricing Studies* 1 (1): 74–95.

Fett, Nicholas, and Richard Haynes. 2017. “Liquidity in Select Futures Markets.” *White Paper. Office of the Chief Economist, United States Commodity Futures Trading Commission*. https://www.cftc.gov/sites/default/files/idc/groups/public/@economicanalysis/documents/file/oce_liquidityfuturesmarkets.pdf.

Fock, Albert. 2024. “Trade Size and Transaction Costs in Over-The-Counter Markets.” Master’s Thesis, Universidade NOVA de Lisboa (Portugal). https://search.proquest.com/openview/fa910bd6b945b267c5d8f6aec4657d17/1?pq-origsite=gscholar&cbl=2026366&diss=y.

French, Kenneth R. 1980. “Stock Returns and the Weekend Effect.” *Journal of Financial Economics* 8 (1): 55–69.

Gai, Prasanna, and Sujit Kapadia. 2010. “Contagion in Financial Networks.” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466 (2120): 2401–23. https://doi.org/10.1098/rspa.2009.0410.

Golan, Amos. 2007. "Information and Entropy Econometrics — A Review and Synthesis." *Foundations and Trends® in Econometrics* 2 (1–2): 1–145. https://doi.org/10.1561/0800000004.
———. 2017. *Foundations of Info-Metrics: Modeling and Inference with Imperfect Information*. New York, NY: Oxford University Press.
Golan, Amos, George G. Judge, and Douglas Miller. 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. Series in Financial Economics and Quantitative Analysis. Chichester [England] ; New York: Wiley.
Golan, Amos, George Judge, and Jeffrey M. Perloff. 1996. "A Maximum Entropy Approach to Recovering Information From Multinomial Response Data." *Journal of the American Statistical Association* 91 (434): 841–53. https://doi.org/10.2307/2291679.
Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 254–77. https://doi.org/10.1016/j.jeconom.2021.03.014.
Greenwood, Robin, Augustin Landier, and David Thesmar. 2015. "Vulnerable Banks." *Journal of Financial Economics* 115 (3): 471–85. https://doi.org/10.1016/j.jfineco.2014.11.006.
Halkiewicz, Stanisław. 2024. "What Are Difference-in-Difference Models? An Overview of History, Theory and Applications." In , 219–31.
Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66 (5): 1017–98. https://doi.org/10.2307/2999630.
Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies* 64 (4): 605–54. https://doi.org/10.2307/2971733.
Hull, John. 2022. *Options, Futures, and Other Derivatives*. Eleventh edition, Global edition. Harlow, England: Pearson4060 1 Online-Ressource (880 Seiten).
Jackson, Matthew O., and Agathe Pernoud. 2021. "Systemic Risk in Financial Networks: A Survey." *Annual Review of Economics* 13 (1): 171–202. https://doi.org/10.1146/annurev-economics-083120-111540.
Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–48. https://doi.org/10.1086/262072.
Kleist, Karsten von, and C. Mallo. 2011. "OTC Derivatives Market Activity in the Second Half of 2010." Bank of International Settlements. https://www.bis.org/publ/otc_hy1105.pdf.
LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *The American Economic Review* 76 (4): 604–20.
Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355. https://doi.org/10.1257/jel.48.2.281.
Lester, Richard A. 1946. "Shortcomings of Marginal Analysis for Wage-Employment Problems." *The American Economic Review* 36 (1): 63–82.
Liu, Jun, Francis A. Longstaff, and Ravit E. Mandell. 2006. "The Market Price of Risk in Interest Rate Swaps: The Roles of Default and Liquidity Risks." *The Journal of Business* 79 (5): 2337–59. https://doi.org/10.1086/505237.
Loon, Yee Cheng, and Zhaodong Ken Zhong. 2014. "The Impact of Central Clearing on Counterparty Risk, Liquidity, and Trading: Evidence from the Credit Default Swap Market." *Journal of Financial Economics* 112 (1): 91–115. https://doi.org/10.1016/j.jfineco.2013.12.001.

Loon, Yee Cheng, and Zhaodong (Ken) Zhong. 2016. “Does Dodd-Frank Affect OTC Transaction Costs and Liquidity? Evidence from Real-Time CDS Trade Reports.” *Journal of Financial Economics* 119 (3): 645–72. https://doi.org/10.1016/j.jfineco.2016.01.019.

Mateus, Cesario, and Fabio Faltoni. 2015. “How Swap Execution Facilities Will Reshape the OTC Derivative Swap Market?” *Journal of Stock & Forex Trading* 4 (3). http://gala.gre.ac.uk/id/eprint/15000/.

McAlley, Eric. 2024. “Examining the Impact of Central Clearing and Swap Execution Facilities on Interest Rate Swap Spreads and Their Determinants.” *McAlley, E.(2024). Examining the Impact of Central Clearing and Swap Execution Facilities on Interest Rate Swap Spreads and Their Determinants. Risk Governance and Control: Financial Markets & Institutions* 14 (1): 8–18.

McPartland, John W. 2009. “Clearing and Settlement of Exchange Traded Derivatives.” *Chicago Fed Letter* 267 (2). https://fraser.stlouisfed.org/files/docs/historical/frbchi/policydiscussion/frbchi_policy_2009-02.pdf.

Menkveld, Albert J., and Guillaume Vuillemey. 2021. “The Economics of Central Clearing.” *Annual Review of Financial Economics* 13 (1): annurev-financial-100520-100321. https://doi.org/10.1146/annurev-financial-100520-100321.

Minton, Bernadette A. 1997. “An Empirical Examination of Basic Valuation Models for Plain Vanilla U.S. Interest Rate Swaps.” *Journal of Financial Economics* 44 (2): 251–77. https://doi.org/10.1016/S0304-405X(97)00005-6.

“OTC Derivatives Statistics at End-June 2024.” 2024. Bank of International Settlement.

Pirrong, Craig. 2011. “The Economics of Central Clearing: Theory and Practice.”

Rambachan, Ashesh, and Jonathan Roth. 2023. “A More Credible Approach to Parallel Trends.” *The Review of Economic Studies* 90 (5): 2555–91. https://doi.org/10.1093/restud/rdad018.

Randall, Oliver. 2015. “Pricing and Liquidity in Over-the-Counter Markets.” *Available at SSRN 2590351*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2590351.

Rochet, Jean-Charles, and Jean Tirole. 1996. “Interbank Lending and Systemic Risk.” *Journal of Money, Credit and Banking* 28 (4): 733–62. https://doi.org/10.2307/2077918.

Roll, Richard. 1984. “A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market.” *The Journal of Finance* 39 (4): 1127–39. https://doi.org/10.1111/j.1540-6261.1984.tb03897.x.

Roth, Jonathan. 2022. “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends.” *American Economic Review: Insights* 4 (3): 305–22.

Rubin, Donald B. 1976. “Inference and Missing Data.” *Biometrika* 63 (3): 581–92. https://doi.org/10.1093/biomet/63.3.581.

Sant’Anna, Pedro H. C., and Jun Zhao. 2020. “Doubly Robust Difference-in-Differences Estimators.” *Journal of Econometrics* 219 (1): 101–22. https://doi.org/10.1016/j.jeconom.2020.06.003.

“SDR View, Capped Notional Changes.” 2013. *Clarus* (blog). August 2013. https://www.clarusft.com/sdr-view-capped-notional-changes/.

Skarr, Doug, and Kristin Szakaly-Moore. 2007. “Understanding Interest Rate Swap Math & Pricing.” California Debt and Investment Advisory Commission. https://www.treasurer.ca.gov/cdiac/publications/math.pdf.

Smith, Clifford W., Charles W. Smithson, and Lee Macdonald Wakeman. 1988. "The Market for Interest Rate Swaps." *Financial Management* 17 (4): 34–44. https://doi.org/10.2307/3665765.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 175–99. https://doi.org/10.1016/j.jeconom.2020.09.006.

Sun, Tong-sheng, Suresh Sundaresan, and Ching Wang. 1993. "Interest Rate Swaps: An Empirical Investigation." *Journal of Financial Economics* 34 (1): 77–99. https://doi.org/10.1016/0304-405X(93)90041-9.

Swap Data Recordkeeping and Reporting Requirements. 2012. *CFR*. Vol. 17 CFR.

Tarbert, Heath P., and Daniel J. Grimm. 2021. "The CFTC's Swap Data Overhaul." *Florida State University Business Review* 20:1.