The Labor Hub

Job Market Paper Series No. 14

Advisor: David Margolis

DRUG LORDS IN STEM

Maria Montoya-Aguirre

Drug Lords in STEM

Maria Montoya-Aguirre*

Job Market Paper

This version: October 30, 2026

Click here for the most recent version.

Abstract

Education is well established as a protective factor against criminal involvement because it improves legal labor-market opportunities. But what happens when illegal markets demand skilled labor? I study how a boom in chemistry-intensive illegal drug production reshaped human capital investment in Mexico. China's 2019 ban on fentanyl pushed its synthesis into clandestine Mexican laboratories, expanding synthetic drug production. Using a difference-in-differences design, I compare commuting zones with pre-existing synthetic-drug infrastructure to those without before and after the ban. Chemical byproducts of drug synthesis are more often detected in the water of these exposed areas than elsewhere, indicating they are production sites. After the ban, school dropout fell in exposed zones, while time spent studying and the share of students reaching the final year of high school increased. University enrollment also reallocated toward chemistry-related STEM programs, with no detectable change in total enrollment. Illegal production also raised the legal demand for chemistry skills: the wage premium on chemistry-related degrees and employment in chemistry occupations both rose. Meanwhile, among adults whose schooling predates the ban, fewer chemistry degree holders appear in the survey, suggesting that skilled chemists are entering illegal production. As criminal organizations come to rely on skills that legal employers also value, these findings raise the question of how effective education will be at reducing crime.

JEL Codes: I26, J24, K42, O15

Keywords: human capital, education, organized crime, drug trafficking, skill-biased technical change

*Paris School of Economics and Université Paris 1 Panthéon-Sorbonne maria.montoya@psemail.eu. I am very grateful for the generous guidance and support from David Margolis, Oliver vanden Eynde, Nina Guyon and Maria Micaela Sviatschi. Thank you to seminar participants at CEPR Development Economics Symposium, Chicago/LSE Conference on the Economics of Crime and Justice, Paris School of Economics, Texas Economics of Crime Workshop, WEC Jr. at the Institute for Fiscal Studies, ViCE seminar, EEA-ESEM conference, EALE conference, and Princeton Development Economics Student Workshop. This work benefited from the EUR grant ANR-17-EURE-0001 and a Horowitz Foundation Dissertation Award.

1 Introduction

Education is widely viewed as a protective factor against criminal involvement. The canonical framework in the economics of crime, originating with Becker (1968), holds that education raises the opportunity cost of crime by improving legal labor market prospects, and a large body of empirical work supports this relationship (Lochner and Moretti, 2004; Machin et al., 2011; Machin and Sandi, 2025). Yet the framework rests on an implicit assumption: that criminal activity primarily demands low-skilled labor, for which education confers no productive advantage. This fits street crime and much of the physical work of the drug trade—cultivating crops, moving contraband, providing muscle for territorial control. But it need not hold when criminal organizations modernize their operations and begin to demand skilled workers. This paper examines the effects of a skill-biased shift in the technology of an illegal market on the human-capital decisions of the youth exposed to it.

I study a setting where technological change in an illegal market clearly raised the demand for a specific skill. Over the past two decades, Mexican trafficking organizations have shifted from plant-based drugs—marijuana, heroin from opium poppy—toward synthetic drugs, particularly methamphetamine and fentanyl. This shift transforms the skills that criminal production rewards: growing cannabis or poppy requires land, agricultural knowledge, and physical labor, whereas synthesizing methamphetamine or fentanyl requires laboratory equipment, chemical expertise, and the ability to follow—or invent—complex production protocols. This demand for chemistry skill is not merely theoretical. Prieto-Curiel et al. (2023) estimate that Mexican criminal organizations are among the country's five largest employers. Leaked documents from Mexico's military intelligence reveal that trafficking organizations actively recruit chemistry professors from universities, and chemistry students and graduates report recruitment attempts (Garcia, 2023; Dudley, 2023).

To identify the effect of exposure to high-skill criminal opportunities, I exploit a policy shock that accelerated the transition towards illegal synthetic drug production. On April 1, 2019, China announced a nationwide ban on fentanyl and its analogues, cutting off the finished product that Chinese firms had been shipping directly to American consumers. Demand did not disappear, so production moved elsewhere. Mexican organizations, already operating methamphetamine laboratories, started fentanyl synthesis in-house, expanding synthetic drug production where they had the necessary infrastructure. Using a difference-in-differences design, I compare commuting zones that had clandestine methamphetamine laboratories before the fentanyl crisis—indicating pre-existing synthetic-drug infrastructure—to those that did not, before and after the ban. The identifying logic is that existing infrastructure gave these zones a comparative advantage in pivoting to fentanyl. This is validated by two facts: post-2019 fentanyl and laboratory seizures concentrate disproportionately

in the pre-exposed zones, and independent water-quality monitoring shows those zones exhibit the chemical signature of methamphetamine synthesis—elevated mercury and toluene—confirming that the exposure measure captures genuine production rather than just differential policing of drug laboratories.

A simple model of occupational choice organizes the consequences of this shock for human capital accumulation. I treat chemistry skill as priced in a local labor market in which a legal and an illegal, drug-producing sector bid for the same, slow-to-replenish pool of chemists, and in which illegal production potentially demands legal inputs—precursors, equipment, and laboratory and commercial services. A rise in the illegal demand for chemistry skill also raises its return in the legal market because of poaching as well as through potential input linkages between the legal and illegal sectors. Youth with different propensities for illegal work choose education, field, and sector. The model predicts that educational investment rises in exposed areas, specifically in chemistry-intensive fields. Additionally, the framework makes it possible to use the estimated legal wage response to gauge the otherwise-unobservable illegal demand shift and corresponding rise in the illegal chemistry premium. Since the shock meets a short-run supply of chemists that cannot be replenished quickly, even a relatively small illegal sector can move chemists' local wages and schooling decisions.

The analysis draws on several data sources and spans 2014 to 2024, a symmetric window around the April 2019 ban. Individual-level educational and labor-market outcomes—enrollment, time use, wages, occupation, and educational attainment for school-age youth and prime-age adults—come from Mexico's National Survey of Occupation and Employment (ENOE), the country's principal labor-force survey. I also measure grade progression up to high school using *Formato 911*, the administrative census of the Mexican school system. Tertiary enrollment comes from administrative data of the National Association of Universities and Higher Education Institutions (ANUIES). I allocate tertiary education enrollment to students' origin commuting zones using student-migration patterns from the 2010 Population Census and classify programs by chemistry intensity from detailed curriculum analysis of programs at Mexico's largest public university. Treatment is built from military records of clandestine-laboratory seizures maintained by Mexico's Secretariat of National Defense (SEDENA); national water-quality monitoring (RENAMECA) is used to validate measures of synthetic drug production. I construct commuting zones to use the local labor market as the unit of analysis, and study the 1,226 commuting zones observed in the labor force survey.

The results overturn the conventional wisdom about education and crime. Following China's ban, exposure to synthetic-drug production *reduced* school dropout. Middle-school dropout fell by 1.2 percentage points (a 21% reduction relative to the control mean of 5.7%) and high-school

dropout by 2.3 percentage points (a 9.2% reduction relative to the control mean of 24.8%). Event-study estimates show no visual evidence of pre-trends; the effects emerge after 2019, grow thereafter, and are largest at the end of the academic year. Youth in exposed areas spent more time studying—roughly 3% more at middle-school age and 5% at high-school age—and 4–5% less on household chores, even during the school closures when formal enrollment was administratively frozen. Administrative grade-progression records confirm that the contemporaneous reduction in dropout translates into higher educational attainment. The cohort in the second-to-last grade of high school when the ban hit—which crossed into the final grade in September 2019, after the ban but before the March 2020 school closures—was 6.2 percentage points more likely to reach the final grade in exposed zones (an 11% increase), an *attainment* effect measured free of pandemic confounding. The other cohorts, which crossed into the final grade after the closures, also show a positive, though smaller, effect.

More tellingly, the education response is skill-specific. Enrollment in STEM and health programs with chemistry content rose by 5.9% in exposed areas, while non-chemistry STEM fields showed no significant change and overall tertiary enrollment was unchanged. School attendance at university age in the labor force survey did not move either, so the shock did not alter the number of young people who go to university. This compositional pattern—a shift toward chemistry-intensive fields with no change in the total—is the distinctive signature of a *skill-specific* shock: a general rise in the returns to schooling would lift enrollment across fields, whereas a chemistry-specific shock reallocates students toward chemistry. The effect grows over the event-study horizon as cohorts exposed during secondary school reach university age. Decomposing by gender, women reallocate *within* STEM—from non-chemistry to chemistry-intensive programs—while men's increased chemistry enrollment appears to draw from non-technical fields, though these gender-specific estimates are less precise.

I show that the demand for chemistry skills rose across the exposed labor markets, suggesting, through the lens of the model, spillovers of the shock to the legal sector through both poaching and input linkages. Exposure to synthetic drug production raised both the price and the quantity of chemistry skill in the legal labor market. The wage premium on a chemistry-relevant degree rose by 6% and employment in chemistry occupations by about 19%. These effects are specific to chemistry: the change in the return to non-chemistry STEM degrees is half as large but statistically insignificant, and the overall wage and employment effects are flat. Alternative explanations cannot account for this pattern: a supply inflow of chemists would raise employment while lowering the wage; poaching of scarce chemists would raise the wage while cutting employment; a general violence- or displacement-based account would move general quantities without raising the price of

the specific skill. In sum, the illegal shock reaches all youth in the area instead of just crime-bound ones because the legal and illegal sectors hire from the same local pool of chemists and the illegal production requires inputs from the legal firms.

I find suggestive evidence consistent with skilled chemists being drawn into illegal production. Among workers who completed schooling before the shock, the observed stock of chemistry degree holders in the labor force survey fell by 4.6% in exposed zones, with placebo fields flat, even as observed employment in chemistry occupations rose. These missing chemists could have left exposed zones, for instance to avoid recruitment, or remained in the area moved into the illegal sector, and concealed their activity by declining to answer the survey. I cannot fully separate the two explanations, but several patterns are more consistent with entry into illegal production. I find no evidence of general out-migration: homicides and other crimes did not rise with exposure, the legal return to chemistry rose within these zones, the out-of-state share of the workforce fell slightly rather than rising, and non-interviews from vacated dwellings did not increase. Still, these tests cannot rule out selective out-migration of chemists. At the same time, household refusals to answer the labor survey rose by 0.7 percentage points in exposed zones. This rise is larger than what the missing chemists alone could produce, likely reflecting broader avoidance of the survey with increased exposure to illegal production. It does not seem to be driven by fear of victimization: extortion reports did not rise, and there is no change in item non-response or income heaping, which would more clearly signal concealment of information out of fear of being targeted.

I observe neither the illegal chemistry premium nor the size of the illegal demand shock. However, since the model is invertible, I can estimate what magnitude of illegal activity would be required to produce the observed rise in the legal chemistry premium—and find that it need not be large. Reproducing that rise requires the illegal chemistry premium to have risen by about a quarter—between 14% and 61% across plausible parameters—and illegal demand for chemistry skill to have grown several-fold to roughly thirtyfold over 2019–2024; a sector employing only a few percent of local chemists produces effects of this size, because its short-run supply is nearly fixed. Independent evidence suggests that the illegal sector is large enough to explain these findings.

I subject the findings to extensive robustness checks, the most important concerning how exposure is measured. A laboratory seizure records production and its *detection* at once: where enforcement is weak, production goes unrecorded, so a seizure-based treatment risks measuring the geography of policing rather than of production, a source of measurement error. I address this in two ways. First, I measure where synthesis is *viable* rather than where it is caught, training Random Forest, XGBoost, and ensemble models on pre-shock fundamentals—proximity to the U.S. border and to the Pacific ports through which precursor chemicals arrive, road connectivity, the

local presence of chemical firms and university chemistry programs, terrain, and cartel presence—while excluding measures of police presence and state capacity. Spatial cross-validation predicts each region from geographically distant ones, which limits how much the models can learn from local policing patterns, and the estimates hold under augmented exposure measures that add the predicted zones to the seizure-based ones. Second, I validate exposure measures against a chemical trace of production that does not depend on policing to be observed: methamphetamine synthesis by the P2P route uses mercury as a catalyst and toluene as a solvent, both of which discharge into local waterways. I find that in national water-quality monitoring both the seizure-based and the predicted exposure measures predict elevated concentrations of exactly these two compounds. The high-school dropout and attainment results also hold under alternative estimators that address identification concerns—doubly robust and synthetic difference-in-differences—and the attainment result is robust to moderate violations of parallel trends.

I also rule out contemporaneous shocks as alternative explanations for my findings. Treated areas were slightly *less* exposed to Mexico's 2019 minimum-wage increases, and within control areas minimum-wage exposure *reduced* dropout—the opposite of what would be needed to explain my results. COVID-19 school closures were mandated uniformly at the national level, and effects on attainment emerge before the closures. Finally, the 2019 rollback of the conditional cash-transfer program Prospera raised dropout in high-penetration areas (Marquez-Padilla et al., 2025), yet the treatment effect survives controlling for it.

The results of this paper carry important policy implications. When criminal organizations also reward the skills that schools teach, education will be less effective at making legal alternatives more attractive than criminal ones, and its protective effect may be limited. The result is increasingly relevant because criminal organizations in many regions of the world are becoming transnational enterprises that rely on skills legal employers also value: chemistry to synthesize drugs, financial and legal expertise to launder their proceeds, and information technology to run online fraud.[1] Where criminal and legal employers compete for the same skills, expanding education and access to good legal jobs also expands the pool of skilled workers that criminal organizations can recruit. This also relates to qualitative evidence in other Latin American settings that points to criminal organizations recruiting through schools and using their infrastructure for criminal activity. Crime prevention in these settings has to rely more on other tools, such as enforcement against criminal

[1]Moving the proceeds of crime through the financial system requires accounting, legal, and financial expertise (United Nations Office on Drugs and Crime, 2023). Losses from online fraud reported to the FBI reached $20.9 billion in 2025, five times their level in 2020, and scam centers, increasingly located across various developing countries, recruit workers with information technology and language skills (Federal Bureau of Investigation, 2025; INTERPOL, 2023,0). European law enforcement describes organized crime as increasingly dependent on technical expertise and digital infrastructure (Europol, ed, 2025).

and legal employers that support the drug trade, as well as the social and cultural norms that can make criminal work less attractive.

This paper contributes to three literatures. First, it contributes to the economics of crime and, in particular, to the study of how education and crime interact. In the canonical framework, education reduces crime because it raises the return to legal work and therefore the opportunity cost of criminal activity (Becker, 1968; Ehrlich, 1973). A large empirical literature supports this prediction: more schooling reduces arrests, incarceration, and both violent and property crime (Lochner and Moretti, 2004; Machin et al., 2011; Bell et al., 2022; Machin and Sandi, 2025), and lower legal wages and adverse labor market shocks increase criminal participation (Grogger, 1998; Dube and Vargas, 2013; Dix-Carneiro et al., 2018). A second strand studies the reverse direction of causality and finds that criminal activity lowers educational investment. Drug trafficking raises school dropout and youth employment in illegal activities (Sviatschi, 2022), gangs restrict the mobility and opportunities of young people (Melnikov et al., 2025), and local violence lowers student achievement (Monteiro and Rocha, 2017; Chang and Padilla-Romo, 2023). In both strands, criminal activity rewards physical labor and, implicitly, few skills that can be taught in formal education. The one exception I am aware of, Campaniello et al. (2016), shows that schooling is associated with higher earnings among members of the Italian-American mafia engaged in skill-intensive activities. I provide causal evidence that the effect of criminal opportunities on education depends on the skills that criminal production requires. When an illegal market adopts a technology that rewards a specific skill, criminal opportunities can increase educational investment, particularly in the fields that teach it. I also show that this happens partly because the legal sector increasingly rewards the same skills as the criminal sector.

Second, the paper contributes to the literature on drug control policy and the study of organized crime as an industry. Most of this work studies how enforcement affects violence and the organization of trafficking. Interdiction and the capture of leaders increase violence and disrupt supply only temporarily, as criminal organizations adapt by reorganizing, changing routes, and moving production (Dell, 2015; Castillo et al., 2020; Castillo and Kronick, 2020; Lindo and Padilla-Romo, 2018; Mejía and Restrepo, 2016). Supply-side drug policies in the United States had similar unintended effects, moving consumers from one drug to another (Alpert et al., 2018; Evans et al., 2019). A broader literature documents the costs of organized crime for migration (Daniele et al., 2023; Aldeco Leo et al., 2024), politics (Alesina et al., 2019), local economic activity and firms (Pinotti, 2015; Mirenda et al., 2022), and the governance of communities (Blattman et al., 2025; Melnikov et al., 2025). I show that a supply-side policy can also change the production technology of criminal organizations and the skills they demand. China's ban on fentanyl moved the synthesis

of the drug to Mexico, where it raised the demand for chemistry skills in the areas with existing production capacity. This new technology also relies on the legal sector.

Existing work shows that organized crime infiltrates legal firms to launder money and to gain economic and political influence (Mirenda et al., 2022). In this setting, legal firms also support production itself: front companies in the chemical and pharmaceutical sectors supply precursor chemicals (U.S. Department of the Treasury, 2025), and registered firms can source inputs for clandestine laboratories and enroll their workers in the social security system (Section 2). Work on organized crime and firms has mostly studied the costs that crime imposes on legal businesses. Here, criminal production depends on legal firms as suppliers and employers, so its growth can also expand legal activity. This blurs the distinction between the legal and illegal sectors, since the same firm, input market, and pool of workers can serve both. Through these input links and through competition for the same workers, the change in illegal production raised the returns to chemistry skills in the legal labor market and changed the schooling and field choices of young people in the producing regions. Drug policy can therefore shape human capital investment in producing countries, an effect that may outlast short-lived enforcement episodes.

Finally, the paper relates to the literature on how labor demand shapes investment in education. Local legal labor demand shocks change how much schooling young people acquire (Atkin, 2016; Shah and Steinberg, 2017; Charles et al., 2018; Adukia et al., 2020). Labor market conditions also shape what students study: expected earnings and local industry composition affect the choice of university and major (Delavande and Zafar, 2019; Weinstein, 2022; Blom et al., 2021). This paper shows that the returns to criminal activity can shape educational investment in the same way. The shocks studied in this literature typically raise the demand for broad groups of workers, and the evidence on field choice relies on broad differences in local industry composition or on business cycles. In this case, the demand rises for one clearly identified skill, chemistry, which maps directly to specific fields of study through the courses they offer. This allows me to show that the effect is concentrated in fields with chemistry courses, with no change in other fields even within STEM. More broadly, this paper highlights that when multiple sectors draw from the same pool of specialized workers, a demand shock in one sector can affect education through the equilibrium responses in the other, in this case through labor poaching and input linkages. This channel is especially relevant in markets for specific skills, where the supply of trained workers responds slowly and wages adjust first (Freeman, 1976; Ryoo and Rosen, 2004; Bound et al., 2015). In my setting, it explains why the schooling response extends to young people who would never work in the illegal sector.

The remainder of this paper proceeds as follows. Section 2 provides background on drug traf-

ficking in Mexico, the drivers of the transition to synthetic drug production, and its labor market implications. Section 3 develops the theoretical framework—a local market for chemistry skill—that generates the paper's testable predictions and that can be inverted to gauge the size of the illegal demand shift. Section 4 describes the data sources and construction of key variables, and Section 5 presents the empirical strategy. Section 6 reports the main results on dropout, grade progression, and tertiary enrollment. Section 7 presents robustness checks addressing alternative treatment definitions, alternative estimators, and contemporaneous shocks. Section 8 turns to the labor-market mechanism: it establishes the low-violence, low-enforcement environment, documents the rise in the price and quantity of chemistry skill in the legal market, and discusses the disappearance of credentialed chemists from observation. Section 9 uses the model's invertibility to gauge the size of the illicit sector and benchmark it against external evidence, and Section 10 concludes.

2 Context

2.1 Mexico's Drug Trafficking Industry

Mexico is the primary source and transit country of illegal drugs entering the United States, the world's largest consumer market. The scale of this industry is enormous: the United Nations Office on Drugs and Crime estimates that inward illicit financial flows from trafficking heroin, cocaine, and methamphetamine averaged roughly $12 billion per year over 2015–2018—comparable to the value of Mexico's agricultural exports over the same period—while U.S. and Mexican government estimates place the drug-trafficking proceeds flowing from the United States into Mexico at $19–29 billion annually (United Nations Office on Drugs and Crime, 2023; U.S. Immigration and Customs Enforcement, 2010).

Major drug trafficking organizations, often called "cartels," control production, transportation, and distribution networks across Mexico and involve sophisticated logistics, financial operations, and systematic corruption of public officials to shield their operations from law enforcement. In 2018, more than two thirds of Mexico's population (71.2%) lived in a municipality where organized crime groups where involved in at least one violent event, about 51% were living where a major cartel is present. Historically, Mexican drug production centered on plant-based drugs like marijuana and heroin. Cannabis and opium poppy cultivation has occurred in Mexico for decades. Both crops are grown primarily in mountainous regions with suitable climate and terrain, particularly in states like Sinaloa, Durango, and Guerrero. In recent years, these organizations have diversified into other criminal activities like extortion, human trafficking, oil theft, and the focus of this paper, synthetic drug production.

2.2 The First Synthetic Drug Expansion: methamphetamine

Beginning in the late 2000s and accelerating through the 2010s, Mexican trafficking organizations increasingly diversified towards synthetic drug production, particularly methamphetamine (Figure 1). This transition was driven by several factors: disruptions to the availability of precursor chemical supplies in the United States that had previously supported domestic meth production; growing demand from American consumers; and the higher profit margins available from synthetic drugs, which require no land and can be produced year-round wherever the necessary inputs can be procured (Asmann, 2021).

Methamphetamine production requires fundamentally different capabilities than agricultural drug production. Synthesizing methamphetamine involves chemical reactions, precursor procurement, laboratory setup and operation, and quality control. While "shake and bake" methods allow small-scale production with minimal expertise,[2] the large-scale operations that supply wholesale markets require genuine chemical knowledge: understanding reaction kinetics, handling volatile and toxic substances, troubleshooting production problems, and adapting methods when precursor availability changes.

Mexican seizure data reflects this expansion. Figure 1a shows that between 2005 and 2015, authorities dismantled an increasing number of clandestine laboratories, predominantly producing methamphetamine. The geographic concentration of these seizures, shown in Figure 1c, reveals that synthetic drug production was not uniformly distributed: only certain regions developed specialized production capabilities, particularly in areas with proximity to precursor chemical supply routes through Pacific coast ports, U.S. border access, and historical presence of criminal organizations.

2.3 The Second Synthetic Drug Expansion: fentanyl

The opioid crisis in the United States created the demand conditions for the drug production transformation studied in this paper. Beginning in the 1990s, aggressive marketing of prescription opioids led to widespread addiction (Van Zee, 2009). When the 2010 reformulation of OxyContin and related regulatory crackdowns restricted access to prescription pills, users turned to illegal substitutes—first heroin (Alpert et al., 2018; Evans et al., 2019), then, around 2015, fentanyl.

[2]The "shake and bake" or "one-pot" method combines pseudoephedrine with volatile reagents—typically lithium stripped from batteries, a source of ammonia such as ammonium nitrate, and a flammable solvent—inside a single sealed container that is shaken to drive the reaction. It requires little equipment and only small quantities of precursor, but the reaction is prone to fire and explosion, has a high failure rate, and yields small batches of low-purity product. It is therefore ill-suited to the sustained, high-volume output required to supply wholesale and export markets, which is why the industrial-scale production discussed here relies on trained personnel rather than amateur "cooks" (National Drug Intelligence Center, 2009).

Fentanyl is a synthetic opioid 50-100 times more potent than morphine that offers enormous advantages to suppliers. It requires no agricultural land, can be produced all year-round production, and its extreme potency means that small shipments yield large quantities of street-level product (National Institute on Drug Abuse, 2021). Chinese legal chemical manufacturers, operating with minimal regulatory oversight, produced fentanyl and intermediaries shipped it directly to American consumers through the postal system, often in packages small enough to evade detection (U.S. Senate Permanent Subcommittee on Investigations, 2018). By 2016, fentanyl had become the leading cause of overdose deaths in the United States (Figure A.1).

On April 1, 2019, responding to sustained diplomatic pressure from the United States, China announced a ban on fentanyl and its analogues (henceforth, the *fentanyl ban*), effective May 1, 2019 (Congressional Research Service, 2019).[3] The ban disrupted supply but did not eliminate American demand. Mexican trafficking organizations, already operating methamphetamine laboratories and controlling distribution routes into the United States, seized the opportunity: China's role shifted from exporting finished fentanyl to supplying precursor chemicals to Mexico, which Mexican organizations now synthesize domestically (U.S. Drug Enforcement Administration, 2020) and then smuggle through the U.S.-Mexico border. This fentanyl expansion built on existing methamphetamine infrastructure—trained personnel, supplier relationships, laboratory equipment, and distribution channels—giving established meth-producing regions a clear advantage in expanding to fentanyl.

Figure 1 illustrates these trends. Panel A shows clandestine laboratory seizures by the Mexican government, which fluctuated between roughly 100 and 250 per year through the 2010s before rising sharply after 2020, to nearly 950 by 2023. Panel B shows fentanyl seizure volumes, which climbed from negligible levels before 2017 to about 340 kilogram-equivalents in 2018 and, after a temporary dip in 2019, grew rapidly to over 1,500 by 2023. Notably, fentanyl did not replace methamphetamine: the two are pharmacologically distinct—one a synthetic opioid, the other a stimulant—but they seem to be complements rather than substitutes, with methamphetamine production continuing to grow alongside fentanyl and the two increasingly consumed together.[4]

[3]Formally, China placed all fentanyl-related substances under national control as a single class ("class scheduling"), rather than banning individually named compounds one by one. This brought every current and future variant of the fentanyl molecule under control at once, closing the loophole by which producers had evaded regulation by making minor chemical modifications to synthesize novel, technically legal analogues.

[4]The share of U.S. overdose deaths involving both fentanyl and a stimulant rose from under 1% in 2010 to roughly 30% by 2021, making stimulants the most common drug class found alongside fentanyl in fatal overdoses (Friedman and Shover, 2023).

2.4 Labor Market Implications

These organizations are not only financially large but also major employers. Combining data on homicides, disappearances, and incarcerations, Prieto-Curiel et al. (2023) estimate that Mexican criminal organizations collectively employ between 160,000 and 185,000 people—placing them, in aggregate, among the country's five largest employers—and must recruit at least 350 new members every week simply to offset losses from arrests and inter-cartel conflict. The scale of this workforce, and the constant recruitment needed to sustain it, means that shifts in what these organizations produce translate directly into shifts in the kind of labor they demand.

Increased synthetic drug production carries at least three labor market implications that are supported by qualitative evidence, which I organize in the model in section 3. The main premise is that this activity creates demand for chemistry skills—individuals with knowledge of chemical reactions, laboratory procedures, and scientific reasoning— that has been documented by various sources. One might ask whether trafficking organizations truly need many trained chemists, or whether a few experts could simply train others on the job. The evidence suggests that skill requirements are substantial and difficult to substitute. Fentanyl synthesis is demanding at every stage—some routes require specific atmospheric conditions, converting pre-precursors into precursors can be fatal if performed incorrectly (Dudley, 2023); the further back along the route producers must work, the more knowledge, equipment and precision are required. Producers are increasingly pushed into more chemically demanding routes as precursor availability is disrupted.[5] What organizations require is therefore a standing capacity to re-engineer production rather than a one-time transfer of a recipe. Untrained "cooks" are costly because accidents destroy product and personnel, while inconsistent dosing attracts law enforcement attention and damages market reputation.[6]

The first consequence is that the shift can raise the returns to chemistry-related skills for individuals willing to work in illegal production. Documents from Mexico's Secretariat of National Defense, leaked in 2022, revealed that the Sinaloa Cartel was actively recruiting chemists.[7] Chemistry professors in Baja California have reported fearing recruitment attempts, with some avoiding

[5] NPP and ANPP, the immediate fentanyl precursors, were added to Table I of the 1988 Convention in 2017; producers shifted routes, prompting the scheduling of 4-AP, 1-boc-4-AP and norfentanyl in 2022 and of 4-piperidone and 1-boc-4-piperidone in 2024. Ten substances were added to Table I in the eight years to 2022, several designer precursors with no known legitimate use, which the INCB attributes to how quickly traffickers substitute uncontrolled chemicals for controlled ones (United Nations Office on Drugs and Crime, 2022; International Narcotics Control Board, 2023).

[6] Early U.S. fentanyl-pill operations blended fentanyl powder with inert fillers in household blenders, distributing the drug unevenly—the "chocolate chip cookie" effect, whereby some pills contain little fentanyl and others a lethal dose. Reporting links this crude mixing, as much as deliberate potency, to many fatal overdoses (Quinones, 2021); DEA testing finds roughly six in ten seized counterfeit pills contain a potentially lethal dose (U.S. Drug Enforcement Administration, 2022).

[7] The documents were part of the "Guacamaya Leaks" hack of SEDENA servers. See Garcia (2023).

mention of their profession to prevent attracting criminal attention (Dudley, 2023).[8]

Second, this demand may spill over to legal chemistry-related sectors through two channels: trafficking organizations compete for chemically skilled workers with legitimate employers—a *poaching* channel—and illicit production draws on *input linkages* to legal firms that supply precursors, equipment, and expertise. Consistent with the input-linkage channel, criminal organizations have established front companies in the chemical and pharmaceutical sectors to launder money and procure precursors.[9] Field interviews suggest these linkages can run deeper still: informants described operators who own clandestine laboratories alongside a registered legal firm that both sources hard-to-obtain inputs for the lab and enrolls the lab's workers in the social security system, so that employees of an illegal operation accrue health insurance and pension contributions.[10] Such arrangements blur the distinction between licit and illicit activity: a single firm can serve both markets, with the infrastructure and expertise built for illegal production now also supporting legal enterprises.

Third, and most consequential for my analysis, the increase of skilled criminal opportunities can shape education choices. On one hand, students interested in the drug trade may invest in the chemistry skills now rewarded by criminal organizations. Consistent with this, a chemistry professor in Sinaloa told InSight Crime that he has had multiple students eager to enter synthetic drug production, some enrolling in his course specifically for that purpose; as their numbers grew, he stopped trying to dissuade them and began teaching laboratory-safety techniques to reduce accidents (Dudley, 2023). On the other hand, the spillovers to the legal sector imply that the effects could extend beyond those interested in a criminal carreer. Because legal and illegal employers bid for chemistry skill from the same local pool, a rise in illegal demand could raise the return to chemistry education for all youth—not only those bound for the illegal sector—who may therefore invest in chemistry skills even if they never intend to work illegally.

3 Theoretical framework

To understand how a shock to the illegal demand for chemistry skill could shape educational choices, I use a Roy-type sorting model of a local labor market following the tradition of modelling crime as

[8] One professor recounted being approached in the 1990s to help produce methamphetamine. In 2021, a chemistry student in Tijuana was killed months before graduation. Prosecutors investigated whether the attack was retaliation for refusing to work for criminal groups. (Garcia, 2023)

[9] In October 2025, the U.S. Treasury sanctioned twelve Mexico-based companies affiliated with the Sinaloa Cartel's Los Chapitos faction, accusing them of using pharmaceutical, laboratory, chemical, cleaning, and real estate businesses to supply precursor chemicals for fentanyl production (U.S. Department of the Treasury, 2025).

[10] Author field interview, Sinaloa, June 2023.

an occupational choices (Becker, 1968; Grogger, 1998). Youth sort across two sectors, a *legal* one and an *illegal*, drug-producing one, and choose whether and in what field to be educated; the return to chemistry skill is set by clearing a local market for chemistry-skilled labor in which both sectors bid. The model generates the predictions that are tested in the rest of the paper. Crucially, it serves to discipline the natural objection that clandestine production is too small a part of the local economy to move wages or schooling by highlighting the role of the legal sector in diffusing the criminal demand for chemistry skills. Even a relatively small criminal sector can influence education choices in this setting. Using the observed wage and educational responses to the shock, the model can then be inverted to gauge the size of the illegal demand shift. The full model, including decision rules, market clearing, existence and uniqueness of equilibrium, and comparative statics, is developed in Appendix C, and its elements are summarized in Appendix Table C.1.

3.1 Setting

Consider the youth of a local labor market, empirically a commuting zone. Each individual makes three choices: whether to complete an additional level of schooling, $e \in \{0, 1\}$, at cost c; if so, a field of study $f \in \{C, N\}$—chemistry-intensive (C) or not (N)—at an idiosyncratic net cost η_i for C; and a sector of employment $s \in \{L, I\}$, legal (L) or illegal (I). Illegal work pays more at the base ($w_{I0} > w_{L0}$) but carries a risk of arrest $p(\theta)$ that is decreasing in an individual's type θ. A higher θ, criminal capital, means a lower chance of being caught—for instance, family or social ties to a trafficking organization that bring local protection, or the know-how to operate without drawing law-enforcement attention.[11] Earnings are

$$w_s(e, f) = w_{s0} + e \cdot \left[\pi_s \, \mathbf{1}\{f = C\} + \beta_s^N \, \mathbf{1}\{f = N\} \right], \qquad s \in \{L, I\}. \tag{1}$$

The two *chemistry premiums* π_L and π_I are prices, determined by clearing a local market for chemistry skill. For simplicity, the non-chemistry returns β_s^N are parameters. Because illegal work carries arrest risk, the illegal sector must pay a compensating differential to attract chemists, so in equilibrium $\pi_I > \pi_L$.

Choices are set by cutoff rules in (θ, η_i). Given schooling and field, the sector choice compares

[11] The arrest-risk interpretation is not essential. The type θ could equally represent an individual's disutility for working in the illegal sector, or any other one-dimensional trait that raises the relative payoff of illegal work. I adopt the arrest-risk formulation because it follows the canonical economics of crime framework in which the risk of detection and associated punishment are the central cost associated to economic activity.

the legal wage with the risk-adjusted illegal payoff,

$$\text{work illegally} \quad \Longleftrightarrow \quad (1 - p(\theta))\, w_I(e, f) - p(\theta)\, P \;>\; w_L(e, f), \tag{2}$$

where P is the penalty on arrest, so lower-risk (high-θ) individuals sort into illegal work; and the chemistry field is chosen when its premium, net of the extra cost η_i, exceeds the non-chemistry return. Both legal and illegal chemistry premiums rise schooling and selection into chemistry fields, but different individuals are affected by the two premiums differently. Those with higher criminal capital θ are more sensitive to the illegal skill premium π_I, while those with lower θ, less likely to join the criminal sector, are more sensitive to the legal premium π_L. The thresholds are solved explicitly in Appendix C.

The legal and illegal sectors are connected by labor poaching and input linkages. They hire chemists from the *same local pool*: a chemist indifferent between them equates risk-adjusted pay at the margin, so the two sectors are gross substitutes in chemistry skill—an illegal boom that bids up π_I pulls marginal chemists across the sorting threshold and, by making them scarce, raises π_L. Illegal production also demands other chemistry-related *inputs*—precursors, equipment, and laboratory and commercial services—so an illegal boom can also shift the *legal* demand for chemistry skill outward. Writing sector demands in iso-elastic form,

$$H_L^d = B\, \pi_L^{-\sigma_L}, \qquad H_I^d = A\, \pi_I^{-\sigma_I}, \tag{3}$$

I summarize the input link by a pass-through $\phi \equiv d \ln B / d \ln A \geq 0$, where $\phi = 0$ is a pure poaching scenario. China's 2019 fentanyl ban, by making domestic synthesis from uncontrolled precursors the profitable production technology, is an outward shift in illegal demand—an increase in A in exposed commuting zones. Market clearing in (π_L, π_I) has a unique solution (Appendix C); the shock's effects follow by comparative statics.

3.2 Predictions

The synthetic drug expansion increases the demand for chemists in the illegal sector, $d \ln A > 0$, raises the equilibrium premiums and increases education outcomes by changing the sorting thresholds.

1. **Both chemistry premiums rise** The illegal premium rises directly; the legal premium rises too, by a smaller amount, $0 < d\pi_L / d\pi_I < 1$. *Empirical counterpart:* the rise in the legal-market wage premium for chemistry-course degrees (Table 4, Panel A).

2. **Legal chemistry *employment* rises only if the supply link is active.** Under pure poaching ($\phi = 0$), legal chemistry *wages* rise while legal chemistry *employment falls*, because the illegal sector draws chemists away along the legal demand curve; legal employment rises if and only if $\phi > 0$. *Empirical counterpart:* the rise in chemistry-occupation employment (Table 4, Panel B).

3. **Educational investment rises.** A higher π_I lowers the schooling threshold for illegal-bound (high-θ) youth—a direct effect—and the induced rise in π_L lowers it for everyone—a spillover—so dropout falls. *Empirical counterpart:* the dropout and enrollment results of Section 6.

4. **Field choice shifts toward chemistry.** If only the *chemistry* premiums rise, the field cutoff moves toward C: chemistry-intensive enrollment rises while non-chemistry and total enrollment do not. *Empirical counterpart:* the compositional tertiary enrollment result (Table 3).

3.3 Estimating the Illegal Demand Shift

Both sectors clear a single local market for chemistry skill, so the legal wage response estimated by the empirical strategy is quantitatively informative about the illegal shock itself. Over the study horizon the education pipeline cannot add chemists—producing one takes years of schooling—so the local stock is fixed on the short term,[12] and every chemist the illegal sector is one the legal sector loses. Moving along the legal demand curve (3), the legal premium then satisfies

$$d \ln \pi_L = \frac{1}{\sigma_L} \frac{dH_I}{H_L}, \tag{4}$$

so the estimated legal premium response equals illegal hiring *as a share of local legal chemistry employment*, scaled by the inverse legal demand elasticity. The effect depends on elasticities rather than by the illegal sector's raw size: with specialized legal demand (σ_L small), an illegal sector that hires only a few percent of local chemists can move the legal premium by several percent. This structure makes the model invertible: the estimated legal premium, considering a range of possible elasticities, recovers the two objects that are not observed in the data—the change in the illegal chemistry premium and the size of the illegal demand shift. I carry out that inversion as a consistency check on the mechanism and confront it with external evidence on laboratory prevalence in Section 9.

[12] While it is possible to recruit chemists from outside of the commuting zone, this paper does not find evidence of increased migration into exposed areas.

4 Data

The analysis combines data from several sources: labor force surveys measuring educational outcomes, administrative records on academic enrollment, military records on drug seizures and laboratory dismantlement, and geographic data for constructing local labor markets. This section describes each data source and the construction of key variables.

4.1 Outcomes

School Dropout (ENOE). The primary measure of educational attainment comes from Mexico's National Survey of Occupation and Employment (ENOE), a quarterly rotating panel that is the country's principal labor force survey. ENOE provides sociodemographic information, including current enrollment for all members of the household, as well as labor market information for members 12 years or older.

I measure dropout by identifying individuals who should be enrolled on a given education level based on their age. Mexico's school entry rules determine that children born before a specific cutoff date should start primary school at age six. Using birth dates, I identify individuals who should be enrolled at each educational level if they progressed on time: primary school age (6–11 years old at the beginning of the school year), middle school age (12–14), and high school age (15–17). Dropout is defined as not currently enrolled in school among individuals of the relevant school age.[13]

The analysis uses data from 2014 to 2024, spanning five years before and after China's fentanyl ban in April 2019. I exclude summer quarters (Q3) due to high seasonality in enrollment—many students report not being enrolled during vacation periods even if they re-enroll in the next academic year, which would confound the dropout measure. Figure B.1 illustrates this pattern. I also exclude the second quarter of 2020, when ENOE temporarily switched to a phone-based survey due to COVID-19 restrictions, resulting in a non-comparable sample.[14] The resulting analysis sample covers 1,263 commuting zones.

School Attainment and Grade Progression (Formato 911). The ENOE measures whether a school-age individual is currently attending school, but does not allow to precisely estimate school attainment and follow grade progression for a given cohort. I use *Formato 911*, the administrative census of the Mexican school system. Collected by the Ministry of Public Education (SEP) at

[13] This approach captures whether school-age individuals are in school, regardless of their actual grade level. A 13-year-old who is enrolled but repeating primary school is not counted as a dropout, while a 13-year-old not enrolled at all is.

[14] INEGI resumed in-person interviews in Q3 2020, though with modified protocols. Results are robust to excluding the entire year of 2020.

the start of every school year, it records enrollment, first-time entrants, and repeaters by grade for the universe of schools in the country. From these counts I construct grade-to-grade continuation rates for each commuting zone and follow successive entry cohorts through the system, measuring the share of students who enter middle school and remain enrolled through the final grade of high school. Because the measure is built from grade-specific enrollment aggregated to the commuting zone, transfers between schools within a zone net out. The data span the 2010–2011 through 2023–2024 school years, cover the same commuting zones as the ENOE sample, and are assembled from two sources over the sample period; I subject their concatenation to extensive harmonization checks.[15] This administrative measure complements the survey: whereas the ENOE captures contemporaneous enrollment at a point in time, Formato 911 captures whether those effects accuumulate and a cohort actually advances to the terminal grade of high school.

University Enrollment (ANUIES). For tertiary (university) enrollment, I use administrative data from the National Association of Universities and Higher Education Institutions (ANUIES), which collects annual enrollment statistics from all Mexican higher education institutions. These data provide enrollment counts by institution, campus, and academic program, but do not include information on students' municipality of origin. This creates a challenge: enrollment is observed at the destination (where universities are located), but treatment is defined at the origin (students' home commuting zones).

To assign enrollment to origin locations, I combine ANUIES data with migration patterns from the 2010 Population Census (10% sample). The census asks respondents where they lived five years earlier, which allows me to construct an origin–destination matrix for individuals of university age (19–24) who attend school. I treat the commuting zone of residence in 2005 as the origin o and the commuting zone of residence in 2010 as the destination d. For each destination, I compute the share of its students who come from each origin, denoted $P_{d,o}$, and allocate enrollment at each destination back to origins using these shares:

$$\tilde{Y}_{o,t} = \sum_{d=1}^{N} P_{d,o} \cdot Y_{d,t} \tag{5}$$

where $Y_{d,t}$ is observed enrollment at destination d in year t, and $\tilde{Y}_{o,t}$ is the imputed enrollment originating from CZ o. Appendix D describes the construction of the matrix in detail.

This procedure measures how tertiary enrollment responds to conditions in students' home com-

[15] Data before 2019 are obtained through the organization Xaber, while the most recent on are downloaded directly from the website of the Ministry of Public Education (Sistema de Información y Gestión Educativa)

muting zones, rather than where universities happen to be located. For commuting zones without a university campus, this is the only way to capture tertiary enrollment responses. For commuting zones with campuses, the measure captures both local students and those who migrate elsewhere for university.

Most students study in their home commuting zone, so the matrix is concentrated on its diagonal. In 2010, 96.1% of students from treated origins and 92.0% of students from control origins studied in their own commuting zone (Appendix Table D.1). Appendix Figure B.2 characterizes the matrix further. Panel (a) shows that the average origin CZ sends students to 5.14 destination CZs, with a probability-weighted mean migration distance of 167 kilometers. Panel (b) shows that migration probabilities are heavily right-skewed: most origin–destination pairs have near-zero flows, while the shares of students who stay approach 100%. Panel (c) shows that, conditional on migration, flows are approximately log-normally distributed, with most probabilities in the 0.01–0.20 range.

The imputation rests on two assumptions, which Appendix D.2 discusses together with the limitations of this approach. First, the share of an origin's students who study in each destination stays as in 2010. Second, within each destination, students in every field come from the same mix of origins. Under these assumptions, the imputation recovers changes in whether young people enroll and in what they study, and it shrinks the estimates toward zero without changing their sign (Appendix D.2). A comparison with the 2020 Census shows that the share of students who study at home and the split between treated and control destinations changed little after the ban. The results do not depend on the matrix: they are nearly identical with separate matrices for women and men, with a matrix built from the 2020 Census, and with enrollment measured where each campus is located (Appendix H.6).

To examine education choices related to the specific skill content of the shock, I classify academic programs by their chemistry content using detailed curriculum analysis. Drawing on publicly available curricula (*mapas curriculares*) from the National Autonomous University of Mexico (UNAM), Mexico's largest and most prestigious public university, I systematically measure the chemistry intensity of different academic programs. I then map all programs in the ANUIES tertiary enrollment database to these UNAM reference programs to assign chemistry intensity measures to each program nationwide.[16]

Using this approach, I classify academic programs into three categories based on their chemistry course requirements: STEM+H programs with chemistry, STEM+H programs without chemistry, and non-STEM+H programs. STEM+H fields include Natural Sciences, Mathematics and

[16] Appendix F describes the full classification methodology in detail.

Statistics; Engineering, Manufacturing and Construction; Agriculture, Forestry, Fisheries and Veterinary; and Health and Welfare, following the International Standard Classification of Education (ISCED) 2013 framework.

Programs classified as STEM+H with chemistry are those that require at least one chemistry course according to their matched UNAM curriculum. These include chemical engineering, pharmaceutical sciences, biotechnology, biochemistry, food science engineering, chemistry, biology, and broad engineering degrees like industrial engineering. Among STEM+H programs, 48.2% require at least one chemistry course, with a mean of 1.96 chemistry courses (ranging from 0 to 27). STEM+H programs without chemistry include science, technology, engineering, and health programs where chemistry is absent from the curriculum—for example, architecture, civil engineering, computer science, mathematics, and mechatronics engineering. Medicine and nursing are classified as STEM+H with chemistry as they require introductory chemistry coursework.[17]

Labor Market Outcomes (ENOE). I also use ENOE data to measure wages and employment of prime-age individuals. In particular, I observe returns to education to university degrees across the field definitions established for examining university enrollment. Also, I look at the share of employment across STEM-related occupations, as well as chemistry-related ones specifically, which include chemists, chemical engineers, pharmacologists, laboratory technicians, and workers in the manuacturing of chemical substances and products.[18]

4.2 Treatment Variables

The empirical strategy compares commuting zones with high versus low exposure to the expansion of synthetic drug production following China's 2019 fentanyl ban. High-exposure areas are those most likely to expand production in response to the ban because they possess the infrastructure, expertise, and supply networks required for synthetic drug manufacturing.

I define exposure based on pre-existing methamphetamine production capacity. As described in Section 2, Mexican trafficking organizations developed domestic methamphetamine production during the first synthetic drug expansion in the 2000s and 2010s. Commuting zones that established meth laboratory operations during this period had established access to the key inputs for fentanyl production: clandestine laboratories, trained chemists and cooks, precursor chemical supply chains, and relationships with trafficking organizations. When China's ban disrupted the fentanyl supply to the U.S. in 2019, these areas had a comparative advantage in pivoting to domestic fentanyl synthesis.

[17] Figure F.1 in the Appendix shows the full distribution of chemistry intensity across programs. Figure F.2 lists the most common programs in each category.

[18] The SINCO codes included are: 2222, 2251, 2438, 2612, 2621, 2813, 2814, 7412.

Clandestine Laboratory Seizures (SEDENA). Data on synthetic drug production infrastructure comes from Mexico's Secretariat of National Defense (SEDENA), which maintains records of military operations including the dismantlement of clandestine drug laboratories. These records cover the period 2000–2023 and include the number of laboratories in each municipality per year.

The main treatment variable is a binary indicator equal to one if any methamphetamine laboratory was seized in the commuting zone before 2015. This cutoff ensures that treatment is determined by infrastructure from the first synthetic drug expansion rather than the post-2019 fentanyl expansion.

Figure 2 maps the spatial distribution of treated and control commuting zones, along with subsequent drug production activity. The map reveals strong geographic persistence in synthetic drug production: post-2019 clandestine laboratory seizures (shown as green diamonds) occur predominantly in areas that already had pre-2015 exposure, with a correlation of 0.42 between the exposure variable and post-2019 lab presence. Fentanyl seizures (purple circles) show a somewhat different spatial pattern, with a correlation of 0.21 with pre-2015 exposure. This partial divergence reflects that fentanyl seizures typically occur later in the distribution chain and closer to the U.S. border, capturing final product seizures rather than production sites. The concentration of both pre- and post- production in western and central Mexico reflects proximity to Pacific ports where precursor chemicals arrive and to northern border regions where synthetic drugs are ultimately trafficked into the United States.

Machine Learning Prediction of Production A seizure-based exposure is an incomplete measure of exposure to production. Seizures capture the intersection of two processes: actual production and detection. For example, laboratories in remote areas or where the army has limited presence may evade detection. This introduces one-sided measurement error in the treatment variable because of potential false negatives —commuting zones with no observed seizures may still have clandestine production but are classified as controls. This can attenuate the estimated coefficients, and since detection might correlate with other local characteristics, it can also introduce bias in the estimated treatment effects.

In order to reduce the incidence of false negatives, I construct alternative treatment measures by predicting production viability using pre-shock features that determine where production is viable. I train Random Forest, XGBoost, and ensemble models on pre-shock features that predict where production is viable. These include economic fundamentals (distance to U.S. border, road network connectivity, proximity to Pacific ports, presence of chemical firms and university chemistry programs), demographics (population structure, schooling levels), geography and climate (bioclimatic

variables, elevation, terrain ruggedness), and indicators of illegal activity (cartel presence and agricultural suitability for drug crops). All features are measured in 2015 or earlier. Since neighbouring commuting zones share both production conditions and detection histories, I estimate the models under a spatial cross-validation scheme: Mexico is partitioned into 25 spatial clusters, which are assigned to five folds, and each zone is fitted by a model trained on the other four folds. This ensures that predictions are extrapolations across regions rather than interpolations within similar areas, preventing that an area is predicted to have a lab simply because it is near another area with a lab.

I define the alternative measures of exposure as $\hat{D}_i^{Aug}(\tau) = \max(D_i, \mathbb{1}[\hat{p}_i > \tau])$, one if either a laboratory was seized in the commuting zone before 2015, or the predicted probability of laboratory presence exceeds a threshold τ. I report results across $\tau \in \{0.2, 0.5\}$, rather than selecting a single value because the usual criteria for optimal threshold selection are not appropriate for this application. Selecting a threshold involves maximizing the agreement between predicted and labeled laboratory presence, but the labels are seizures rather than true production here. So, false positives would be penalized under a goodness-of-fit criterion, reproducing the seizure detection pattern I'm trying to correct.

4.3 Validation of Exposure Measures

To validate that these exposure measures capture methamphetamine production and future synthetic drug activity, I examine whether the exposed areas show elevated concentrations of chemicals associated with drug synthesis in independent environmental monitoring data. I use water quality monitoring data from Mexico's National Water Quality Monitoring Network (RENAMECA). This system collected samples at approximately 5,000 monitoring sites across Mexico's major water bodies, measuring concentrations of heavy metals, organic solvents, and other pollutants. I focus on chemicals associated with methamphetamine synthesis using the P2P method: mercury (used as a catalyst in the aluminum-mercury amalgam reduction step) and toluene (used as a solvent for extraction and purification), this was the dominant synthesis route after Mexico restricted pseudoephedrine in 2008. I aggregate site-level measurements to commuting zones using inverse-distance weights from each monitoring site to the commuting zone centroid, constructing annual pollution indices for the period 2015–2020. Appendix I describes the details of this exercise.

Investigative reporting has documented that clandestine laboratories routinely discharge toxic waste directly into water sources. A journalistic investigation found that laboratories extract water from rivers and irrigation canals for production, then return contaminated wastewater containing acids, solvents, and heavy metals to the same water bodies (Lab, 2025). Journalists and environmental activists in Sinaloa report finding laboratory equipment—trays, containers, gas cylinders—in

rivers, often carried downstream by rainfall from mountain laboratories. If my exposure definition reflects synthetic drug production, areas classified as producers should show elevated concentrations of chemical waste in nearby water, particularly mercury and toluene.

Table 1 reports the results of regressions of future laboratory seizures and pollution indicators on different exposure measures. Each column includes the observed exposure seizure measure and one of the machine-learning prediction classifications simultaneously, so each coefficient reflects the variation specific to that measure. Panel A and B vary in the threshold used for the classification of the predicted exposure, in the first case $\tau = 0.2$ and in the second $\tau = 0.50$.

Both seizure-based and predicted exposure measures explain subsequent synthetic drug production measured by future seizures or relevant chemical byproducts in the water. Columns 1 and 2 report an out-of-sample prediction exercise, where the dependent variable is whether a laboratory was seized in the commuting zone after 2015. Commuting zones with a pre-2015 seizure are 8.8-9.6 percentage points more likely to have a laboratory seized after 2015, on top of this, the additional exposure predicted by the machine-learning model is associated with a 3-17 percentage point increase in the likelihood of future laboratory seizures, depending on the threshold used.

Similarly, commuting zones with historical laboratory seizures show mercury detection rates 10.3-14.8 percentage points higher than control ones, and toluene detection rates 2-3.2 percentage points higher. The machine-learning predictions of exposure explain additional variation in both detection rates, by increasing the likelihood of detecting mercury and toluene by up to 7.6 and 2.2 percentage points, respectively. These regressions control for the chemical-to-biological oxygen demand ratio, which measures overall chemical waste in the water, as well as the number of measurement sites so that the detection of mercury or toluene does not reflect broader differences in chemical pollution across areas or different sampling intensity.

4.4 Commuting Zones

I aggregate the analysis to commuting zones (CZs) rather than using municipalities for several reasons. First, labor markets typically span municipal boundaries, particularly in metropolitan areas. Second, aggregation reduces measurement error in treatment status: a single municipality missing a lab seizure may simply have been lucky to avoid detection, but a commuting zone without any seizures more plausibly lacks production infrastructure. Third, commuting zones provide appropriate geographic units for analyzing educational choices, as students often attend schools outside their municipality of residence but within the same city or metropolitan area.

I construct CZs for Mexico following the methodology of Duranton (2015), originally developed to delineate metropolitan areas in Colombia. Using the 2010 Population Census, which includes in-

formation on both municipality of residence and municipality of work, I construct origin-destination commuting matrices. Mexico's 2,478 municipalities are then aggregated into CZs based on commuting flows, resulting in 1,776 commuting zones, of which 1,226 are covered in the analysis sample.[19]

5 Empirical strategy

The empirical strategy compares educational outcomes in commuting zones with pre-existing synthetic drug production infrastructure (treated) to those without (control), before and after China's 2019 fentanyl ban. I estimate separate specifications for dropout (individual-level) and tertiary enrollment (commuting zone-level).

5.1 Dropout before Tertiary Education

For dropout outcomes, I estimate the following difference-in-differences specification at the individual level:

$$y_{i,c,t} = \beta \cdot Exposure_c \times Post_t + \alpha_t + \delta_q + \gamma_c + X_{i,c,t} + \epsilon_{i,c,t} \tag{6}$$

where $y_{i,c,t}$ is a dropout indicator for individual i in commuting zone c at time t (measured quarterly). $Exposure_c$ is an indicator equal to one if any methamphetamine laboratory was seized in commuting zone c before 2015. $Post_t$ is an indicator for periods after April 2019. The specification includes time period fixed effects (α_t), quarter-of-year fixed effects (δ_q) to account for seasonality in enrollment, and commuting zone fixed effects (γ_c). Individual-level controls $X_{i,c,t}$ include age, gender, and household education of adults. Standard errors are clustered at the commuting zone level.

The coefficient β captures the differential change in dropout rates in treated relative to control commuting zones after 2019. A negative β would indicate that exposure to synthetic drug production infrastructure led to reduced dropout rates, consistent with the theoretical prediction that skill-biased technological change in the illegal sector increases returns to education.

5.2 Tertiary Enrollment

For tertiary enrollment, the outcome is measured at the commuting zone level and includes many zeros, particularly for smaller CZs and specialized fields. I therefore estimate a Poisson pseudo-

[19] The remaining commuting zones were not sampled by ENOE during the analysis period.

maximum likelihood (PPML) model (Silva and Tenreyro, 2006):

$$E[\tilde{Y}_{c,t}] = \exp(\beta \cdot Exposure_c \times Post_t + \gamma_c + \alpha_t) \tag{7}$$

where $\tilde{Y}_{c,t}$ is enrollment in commuting zone c in year t, allocated to origin CZs using migration shares as described in Section 4.1. PPML handles zeros naturally and provides consistent estimates under heteroskedasticity (Silva and Tenreyro, 2006). I estimate this specification separately for total enrollment, chemistry-intensive STEM+H enrollment, and non-chemistry STEM+H enrollment to test whether effects are concentrated in fields relevant to synthetic drug production.

5.3 Reaching the Final Grade of High School

The dropout specification measures contemporaneous enrollment. To measure attainment—whether students actually progress to the end of high school—I use the Formato 911 census to follow cohorts of middle-school entrants. For each commuting zone c and entry cohort τ (students entering the first grade of middle school in year τ), I compute $S_{c\tau}$, the share of the cohort that remains continuously enrolled through the final grade of high school, as the product of grade-to-grade continuation rates along the cohort's trajectory.

Because the fentanyl ban occurred in April 2019, in the middle of the 2018–19 school year, cohorts were exposed to it at different stages of their schooling. I index each cohort by the grade $G_{c\tau}$ it was enrolled in at the ban and estimate the cohort event study

$$S_{c\tau} = \sum_{g \neq 12} \beta_g \, Exposure_c \cdot \mathbb{1}\{G_{c\tau} = g\} + \gamma_c + \delta_g + \varepsilon_{c\tau}, \tag{8}$$

where γ_c are commuting-zone fixed effects and δ_g are grade-at-ban fixed effects; observations are weighted by cohort size and standard errors are clustered at the commuting-zone level. The omitted category is the cohort enrolled in the final grade of high school at the ban—the last cohort to reach the terminal grade before the shock. Coefficients β_g for $g < 12$ measure the effect on cohorts that had not yet reached the final grade when the ban hit, and so completed part of their trajectory under the new regime. The coefficient of central interest is β_{11}, for the cohort enrolled in the second-to-last grade of high school (grade 11, approximately age 16) at the ban: this cohort crossed into the final grade in the first post-ban school year (2019–20), recorded in September 2019—after the April 2019 ban but before the March 2020 COVID-19 school closures—making it the one cohort whose attainment response is observed free of pandemic confounding.

I also report a pooled estimate that replaces the grade-at-ban interactions with a single $Exposure_c \times$

$Post_\tau$ term, where $Post_\tau$ indicates cohorts that reached the final grade in or after 2019. Because every post-ban cohort except the 2019 cohort reached the terminal grade during or after the COVID-19 closures, this pooled estimate blends the disruption-free 2019 effect with COVID-affected cohorts and should be read as a conservative summary. Identification mirrors the dropout specification: absent the ban, cohort progression would have trended in parallel across successive cohorts in treated and control commuting zones, a hypothesis the pre-ban cohorts (enrolled in grades 13–15 at the ban, reaching the final grade in 2015–2017) test directly.

5.4 Identification

The key identifying assumption is that, in the absence of China's fentanyl ban, educational outcomes would have evolved similarly in treated and control commuting zones. This assumption would be violated if pre-existing trends differed across treatment status or if other shocks coincided with the 2019 ban and differentially affected treated areas.

I assess the plausibility of parallel trends in several ways. First, I estimate event-study specifications that allow for dynamic treatment effects and visually inspect pre-trends. Second, I use the machine learning-predicted treatment measures, which address potential selection into observed seizures based on enforcement intensity.

Figure B.3 presents standardized differences in baseline characteristics between treated and control commuting zones. Treated CZs differ significantly on several dimensions. In terms of geography and access, they have higher annual precipitation, greater elevation range, larger total area, and are closer to state capitals and Pacific international ports. They also show higher indicators of illegal activity, including greater cannabis and poppy suitability, more cartel presence, and larger cannabis and poppy seizures. Demographically, treated CZs are more urban (lower rural population share), have better infrastructure (higher sewerage and electricity coverage), and a larger share of working-age population. On education, treated CZs have more higher education institutions, campuses, and programs, though baseline enrollment rates and years of schooling are similar and there are no differences in the share of chemistry-intensive programs. Treated CZs also have higher labor force participation, mostly driven by more female employment, and more formal jobs. Notably, treated and control CZs do not differ in the total number of firms or the overall share of chemical industry firms. They differ only in the share of fentanyl-relevant firms—a narrower category that includes specific chemical manufacturing subsectors (such as pharmaceutical and pesticide production), medical equipment suppliers, and wholesale distributors of laboratory equipment. This pattern suggests that treated areas do not simply have larger or more industrialized economies, but rather possess the specific industrial infrastructure relevant to synthetic drug production.

These differences are consistent with drug production locating in areas with specific geographic advantages (access to ports, suitable terrain for concealment) and existing illegal economy infrastructure. While treated and control CZs differ on levels of many characteristics, the identification strategy relies on parallel trends in outcomes, not balance in levels. A natural concern is that some contemporaneous shock correlated with a single unbalanced characteristic, such as proximity to Pacific ports or cartel presence, could confound the results. The machine learning approach provides a partial response to this concern: the ML-predicted treatment is identified from the joint combination of all characteristics in Figure B.3, not from any single variable. For a confounding shock to explain results under both the baseline seizure measure and the ML measure, it would need to correlate with the specific multivariate pattern that predicts synthetic drug production—a more demanding requirement than correlation with any individual characteristic.

Beyond these observable differences, several contemporaneous policy changes could confound the results if they differentially affected treated commuting zones. These include substantial minimum wage increases beginning in 2019 (with larger increases in the Northern Border Free Zone), the discontinuation of Mexico's conditional cash transfer program Prospera, and the COVID-19 pandemic beginning in 2020. In Section 7, I present tests ruling out these alternative explanations.

6 Results

The theoretical framework in Section 3 predicts that a rise in the return to chemistry skills increases educational investment in exposed areas and shifts field choice toward chemistry-intensive programs. This section tests both predictions, first for high school and then for university.

6.1 High school: dropout and attainment

Dropout. High school dropout falls in exposed commuting zones after the fentanyl ban, with no sign of a differential trend before it. Figure 3, Panel (a), presents the event study for young people of high school age (15–17). Before the ban, the coefficients are close to zero and statistically insignificant, with a mean pre-treatment estimate of 0.005 ($p = 0.48$). After the ban, the mean estimate is -0.015 ($p = 0.045$), and the effect grows over time, reaching about -0.03 five years after the ban.

The difference-in-differences estimates imply a reduction in dropout of about 2 percentage points, or 9% of the control mean. Exposure reduces high school dropout by 2.1 percentage points without controls (Table 2, column 1) and by 2.3 percentage points with controls for gender, age,

and household education (column 3, $p < 0.001$), relative to a control mean of 24.8%. The effect is the same for boys and girls (column 4).

The effect is present throughout the school year and is largest at its end. The Mexican school year runs from August to July, so the fourth quarter (October–December) is the start of the year and the second quarter (April–June) is its end. Estimating the model separately by quarter, exposure reduces dropout by 2.0 percentage points in the fourth quarter, 1.7 in the first, and 3.1 in the second (Table 2, columns 5–7). The end-of-year effect is a reduction of 11.5% relative to the quarter-specific control mean of 26.4%, and its event study shows an effect that grows steadily after the ban (Figure 3, Panel (b)). This pattern suggests that students who would have left school during the year are staying until it ends.

The response is concentrated among young people old enough to leave school for work. At middle school age (12–14), exposure reduces dropout by 1.2 percentage points, or 21% of the control mean of 5.7% (Appendix Figure G.1 and Table G.1). At primary school age (6–11), when children are not yet at that margin, there is no effect (Appendix Figure G.2).

Lower dropout comes with more time spent studying. Time-use data from the same survey show that, in exposed areas, the probability of studying in the reference week rises by 2.7 percentage points at high school age and 1.5 percentage points at middle school age. Study time rises by about 5% and 3%, and time on household chores falls by 4–5%. Study time rises after the ban and stays higher during the school closures, when enrollment was administratively frozen (Appendix Figure G.3 and Table G.2).

Attainment. The lower dropout translates into more students reaching the final grade of high school. The survey records whether a young person attends school at a given date, but not whether they complete it. To measure completion, I use the Formato 911 school census, which follows the progression of each cohort through the grades. Figure 4 presents the cohort event study for the share of a middle-school entry cohort that reaches the final grade of high school (grade 12). Cohorts that reached grade 12 before the ban show no differential trend.

The clearest evidence comes from the first cohort exposed to the ban, which completed its progression before the pandemic. This cohort was in grade 11 when the ban was announced and reached grade 12 in September 2019, before the school closures of March 2020. In exposed areas, it was 6.2 percentage points more likely to reach grade 12 ($p < 0.001$), an increase of 11% relative to the control rate of 54.9%. About half of this effect comes from students progressing on time and half from students who stay enrolled although they are behind (Appendix Figure G.4).

Later cohorts also respond, although by less. These cohorts reached grade 12 during or after

the closures, and their estimates are positive but smaller, so the pooled post-ban estimate is 3.1 percentage points ($p < 0.01$, Appendix Table B.1). The share of each cohort reaching the end of middle school (grade 9) changes little (Appendix Figure G.5), which places the response at the transition through high school.

Together, the survey and administrative data show that exposure reduced dropout and increased the share of students who reach the final grade of high school. Because the attainment result comes from a cohort that reached the final grade before the pandemic, it does not depend on the enrollment rules during the school closures.

6.2 University enrollment

To study university enrollment, I assign students to the commuting zone they come from. University enrollment is recorded at each campus, and I assign it to origin commuting zones using a migration matrix from the 2010 Census (Section 4). This allows me to compare enrollment from exposed and control origins, and to measure changes both in whether young people enroll and in what they study. Appendix D discusses the assumptions behind this assignment. Under them, the estimates are about 10% smaller than the true effects, and their sign does not change.

Total enrollment. The shock did not change the number of young people who go to university. Figure 5 presents event studies by field of study. For total enrollment, the pre-treatment coefficients are positive and decline toward zero before the ban (mean pre-treatment estimate 0.028, $p = 0.02$), so exposed commuting zones were on a relatively declining path before 2019. After the ban, the coefficients are small and statistically insignificant (mean 0.018, $p = 0.42$), and the difference-in-differences estimate is zero (Table 3, column 1). Because of the pre-trend, I complement this result with school attendance at tertiary age in the labor force survey, which does not depend on the migration matrix. Attendance does not change in exposed areas, also when the sample is limited to commuting zones with a university campus (Appendix Table H.7 and Figure H.7).

Field of study. Enrollment in chemistry-intensive programs rises by about 6% in exposed areas, while enrollment in other fields does not change. Chemistry-intensive STEM+H programs include, for example, chemical engineering, pharmacy, and biotechnology. Their enrollment shows no differential trend before the ban (mean pre-treatment estimate -0.006, $p = 0.67$) and rises steadily after it, reaching about 0.09 log points, or 9%, four years later (Figure 5). The difference-in-differences estimate is 0.057 ($p < 0.05$), a 5.9% increase (Table 3, Panel A). Enrollment in STEM+H programs without chemistry and in non-STEM+H programs does not change.

Women and men both move toward chemistry, from different fields. Chemistry-intensive enrollment rises by 6.3% for women and 4.9% for men. For women, the increase comes with a decline in other STEM+H programs (-0.049, $p < 0.10$), which suggests a reallocation within technical fields. For men, other STEM+H enrollment does not change (Table 3, Panels B and C, and Appendix Figure G.7).

These results do not depend on how enrollment is assigned to origins. The estimates are nearly identical with separate migration matrices for women and men, with a matrix built from the 2020 Census, and with enrollment measured where each campus is located, without any assignment (Appendix H.6, Table H.6 and Figure H.6). They also hold when the sample is limited to programs observed in every year and when enrollment is measured per available program, so they do not reflect the opening of new chemistry programs (Appendix Figure G.8).

Taken together, the university results show a change in what young people study. Total enrollment and attendance do not change, while enrollment in chemistry-intensive programs from exposed areas increases by about 6%.

7 Robustness

This section shows that the main results hold with alternative measures of exposure and alternative estimators, and that they are not explained by other policy changes of the same period. The university results also hold with alternative ways of assigning enrollment to commuting zones, which Section 6.2 and Appendix H.6 discuss.

7.1 Measurement of exposure

The results hold when exposure is measured in other ways. The baseline measure misses production that was never detected, so Section 4 constructs two alternatives. The first is stricter and requires more than one laboratory seized before 2015. The second adds to the seizure-based measure the commuting zones where a machine-learning model predicts production, for thresholds between 0.2 and 0.5 (Appendix E). Figure 6 reports the estimates. High school dropout falls by 1.7 to 2.9 percentage points across these measures, compared with 2.3 in the baseline, and middle school dropout by 1.3 to 2.4 percentage points, compared with 1.2. The share of the first exposed cohort reaching grade 12 rises by 5.6 to 6.6 percentage points, compared with 6.2, and enrollment in chemistry-intensive programs by 5 to 10%, compared with 5.9%. Total enrollment and enrollment in other STEM+H programs do not change under any measure.

Measures based only on the machine-learning prediction give smaller estimates, which go to zero at high thresholds (Appendix Figure H.1). At high thresholds the model classifies few commuting zones as exposed, so most commuting zones with seizures end up in the control group. At a threshold of 0.5, for example, the random forest classifies only 34 commuting zones as exposed. The comparison then includes exposed areas among the controls, which pushes the estimates toward zero.

7.2 Alternative estimators

The high school and attainment results also hold with estimators that relax the parallel trends assumption. Appendix Figure H.2 compares the baseline estimates with the doubly robust estimator of Sant'Anna and Zhao (2020), which reweights control commuting zones to match exposed ones on baseline characteristics, such as schooling, demographics, firms, and distances to the U.S. border, ports, and cartel areas. With this estimator, high school dropout falls by 4.0 percentage points, compared with 2.3 in the baseline. The share of the first exposed cohort reaching grade 12 rises by 3.8 percentage points, and the share of all post-ban cohorts by 2.5. The synthetic difference-in-differences estimator, which I can apply to the administrative data because they form a balanced panel, gives 5.1 and 2.8 percentage points for these two outcomes. The middle school dropout estimate, on the other hand, falls to -0.003 and is not significant with the doubly robust estimator.

The attainment effect for the first exposed cohort is also robust to violations of parallel trends. Following Rambachan and Roth (2023), Appendix Figure H.3 reports confidence sets that allow the trends of exposed and control areas to differ after the ban. The effect on the first exposed cohort remains positive unless the post-ban difference in trends is more than about 1.7 times as large as the largest difference before the ban. The average effect over all post-ban cohorts is less robust, and it is no longer significant once the post-ban difference exceeds about a third of the largest pre-ban difference.

7.3 Contemporaneous shocks

Three changes in the same period could also affect schooling: the minimum wage increases that began in 2019, the COVID-19 pandemic, and the end of the Prospera cash transfer program in early 2019. I test them using ENOE dropout because the survey also records household income and youth employment, which allows me to examine how each shock could affect schooling.None of the three explains the results (Appendices H.3–H.5).

- *Minimum wage.* Exposed commuting zones had slightly fewer workers affected by the in-

creases, and in control areas a larger share of affected workers was associated with lower dropout. Any bias from the minimum wage therefore works against the results. The estimates are similar outside the northern border zone, where the increases were largest, and when controlling for household income.

- *COVID-19.* Schools closed at the national level. Exposed and control areas had similar numbers of cases and deaths in 2020, and the estimates are similar when the commuting zones with the most cases or deaths are excluded. The effect on the first exposed cohort reaching grade 12 is measured before the closures.

- *Prospera.* The end of the program raised dropout in commuting zones where many households received it. The effect of exposure remains when I allow for this, and when I exclude the most rural commuting zones, where the program was most common.

8 Mechanisms

This section shows that exposure to synthetic-drug production raised the demand for chemistry skills. Exposure did not change local violence, so the schooling response does not clearly reflect a change in the risk of entering illegal work nor is driven by more peaceful conditions. In the legal labor market, measured in the labor force survey, both the wage premium on chemistry-related degrees and employment in chemistry occupations rose, and the new jobs were filled by less-educated local workers. At the same time, the observed stock of adults with chemistry and STEM degrees fell, and survey refusals rose. I cannot separate whether these credentialed workers left the area or stayed and moved into concealed production, but there is no sign of a general out-migration.

8.1 Violence and enforcement

Exposure did not increase local violence. Figure 7 presents the event study for the homicide rate. Before the ban, treated and control commuting zones follow the same path (mean pre-treatment estimate -0.04, $p = 0.66$). After the ban, the estimates are positive but small and statistically insignificant (mean 0.10, $p = 0.24$), and the difference-in-differences estimate is 0.13 ($p = 0.18$, Appendix Table J.1). The result is similar with census population, with the homicide count, and in the commuting zones that were most violent before the ban, where the estimate is 0.21 ($p = 0.07$). Other crimes, such as retail drug sales, kidnapping, injuries, business robbery, and family violence, also show no change at the ban (Appendix Figure J.2). Extortion is higher after the ban relative to

2018, but it was equally high in 2015 and 2016. The increase comes from a low level in 2017 and 2018, and extortion does not change at the ban.

This matters for two reasons. First, a fall in violence could by itself raise schooling (Jarillo et al., 2016; Brown and Velásquez, 2017), and the data show no such fall. Second, entering illegal production did not become more dangerous. Enforcement against fentanyl production was weak during the study period: laboratories were seized, but few people were detained, and there were no major fentanyl-related arrests until late in the period. The arrests of Ovidio Guzmán in 2019 and 2023 did not stop production (Dell, 2015), the United States sanctioned the Sinaloa Cartel for drug trafficking only in December 2021 (U.S. Department of the Treasury, 2025), and the cartel split in 2024, after the study period (Sobrino, 2019).

8.2 Demand for chemistry skills in the legal labor market

The labor force survey records legal employment, both formal and informal, so it measures the legal side of the market for chemistry skills. I look at wages and employment of prime-age workers (25–54) in Table 4.

The wage return to chemistry-related degrees rose in exposed areas. Panel A estimates how the wage premium on each type of degree changed in exposed commuting zones after the ban. The premium on a STEM degree with chemistry courses rose by 5.9 log points ($p < 0.05$, column 6), from a baseline premium of 52 log points. The premium on STEM degrees without chemistry rose by 2.3 log points, and the premium on non-STEM degrees by 3.2 log points, neither statistically significant. Average wages did not change (column 1). Figure 8, Panel (a), shows the event study for the STEM-degree premium: it is flat before the ban (mean 0.018, $p = 0.33$) and higher after it (mean 0.061, $p = 0.02$).

Employment in chemistry occupations also rose. Panel B estimates the share of employment in each occupation group. The share in chemistry occupations, such as chemists, chemical engineers, and chemical-plant operators, rose by 17.6 log points ($p < 0.01$, column 6), about 19%, and the share in STEM occupations by 7.3 log points ($p < 0.01$, column 4). The overall employment rate and the share in non-STEM professional occupations did not change. The event study in Figure 8, Panel (b), is flat before the ban (mean -0.01, $p = 0.89$) and rises after it (mean 0.17, $p = 0.04$). Appendix Figure G.10 shows the event studies for other STEM and other professional occupations.

Wages and employment rising together point to an increase in the demand for chemistry skills. If more chemists had entered the market with no change in demand, employment would rise but wages would fall. If demand had stayed the same and chemists had left the legal market, wages would rise but employment would fall. Only an increase in demand raises both. In the model of

Section 3, this increase reaches the legal market through competition with the illegal sector for chemists and through the demand for legal inputs.

The new jobs in chemistry occupations went to less-educated local workers, as expected when demand expands. Table G.3, Panel A, shows that average schooling fell by 0.15 years among workers in STEM occupations ($p < 0.01$) and by 0.28 years among workers in chemistry occupations ($p < 0.05$), relative to a control mean of 14.4 years. When an occupation hires more workers, the additional workers usually have less schooling than those already employed, so average schooling falls. Panel B shows that the share of workers born in another state did not change in STEM or chemistry occupations, and fell by 0.9 percentage points among all workers ($p < 0.10$). The new chemistry workers therefore came from the local population.

8.3 Illegal production and missing chemists

The illegal market cannot be observed directly, but several facts suggest that its demand for chemistry skills grew. Laboratory seizures after 2019 are concentrated in the exposed commuting zones (Figure 2), and between the years before and after the ban, laboratory seizures grew about 4.4-fold and fentanyl seizures about 8.7-fold (Section 9). Because enforcement was weak, these numbers understate the growth of production.

The observed number of adults with chemistry and STEM degrees fell in exposed areas. Table G.4 follows a fixed group of adults, those aged 30–54 in 2019, whose education was complete before the ban. In exposed commuting zones, the share of this group observed with a STEM degree fell by 3.7% ($p < 0.10$), and the share with a chemistry-related degree by 4.6% ($p < 0.05$). The share with any tertiary degree and the share with a non-STEM degree did not change. Figure 9, Panel (b), shows the event study for the STEM-degree share. The estimates after the ban are negative and stable (mean -0.073, $p = 0.02$). Before the ban, the mean estimate is small and insignificant (-0.027, $p = 0.28$), although the estimate five years before the ban is also negative. Appendix Figure G.9 shows the same pattern for chemistry-related degrees. Because the education of this group was complete before the ban, the decline must come from people leaving the area or from people no longer being observed in the survey.

At the same time, more households in exposed areas refused to answer the survey. Figure 9, Panel (a), shows the event study for the share of inhabited dwellings whose occupants refused the interview. The estimates are flat before the ban (mean -0.003, $p = 0.31$) and rise after it (mean 0.012, $p = 0.01$). The difference-in-differences estimate is 0.7 percentage points ($p < 0.05$), against a control refusal rate of 2.7% (Appendix Table J.2, Panel A). Refusals come from households that live in the dwelling. Dwellings left empty by people who moved away are recorded separately, so

this increase does not reflect out-migration.

I cannot separate whether the missing chemists left the area or stayed and moved into illegal production. Both explanations are consistent with the data. What the data rule out is a general out-migration from exposed areas. Homicides did not rise, so there was no violence to flee, and the legal return to chemistry rose in the exposed areas themselves. The share of workers born in another state fell slightly, while an out-migration of local workers would have raised it. And the decline is specific to STEM and chemistry degrees, while a general out-migration would also reduce the share with other degrees. Any out-migration is therefore likely limited to the credentialed chemists most exposed to recruitment by criminal organizations.

The rise in refusals is unlikely to reflect fear of crime. Violence did not increase, and households that answered the survey did not become less forthcoming. Missing labor income and missing occupation did not change, and reported incomes became, if anything, less rounded (Appendix Table J.2, Panel B). If households had become afraid of being targeted, they would be more likely to withhold or round their income. The pattern is consistent with households linked to illegal production declining the survey, which matches accounts from ENOE field staff at INEGI of recurring difficulty completing interviews in areas of known illegal production.

One implication applies to the results in the previous subsection. The workers who left the survey are likely those with the highest returns to chemistry skills, so the measured increase in the legal wage premium is likely a lower bound.

9 Size of the illicit sector

The mechanism requires an illegal sector large enough to move the legal market for chemistry skills, and this section shows that a small sector is enough. The difference-in-differences design measures the rise in the legal chemistry premium, but not the rise in the illegal premium or the size of the illegal demand shock. Because both sectors hire chemists from the same local pool, the model of Section 3 can be used in reverse: given the estimated rise in the legal premium, the market-clearing conditions tell us how large the illegal premium and the illegal demand shock must have been. This exercise checks whether the mechanism is plausible. It does not identify the effects, which come from the difference-in-differences estimates.

The answer depends on two quantities that the data do not pin down, so I report the results as a function of them. The first is the reallocation elasticity κ_L, which measures how easily chemists move into legal chemistry jobs, from illegal work or from other legal jobs. The second is the relative size of the illegal sector r, the number of chemists working in illegal production as a share of those

in legal chemistry jobs. Figure 10 shows the implied magnitudes over plausible ranges of both. Explaining the estimated legal premium requires the illegal premium to rise by about a quarter and illegal demand for chemistry skills to grow between several-fold and thirtyfold, depending on the size of the illegal sector. Laboratory and fentanyl seizures grew 4.4-fold and 8.7-fold over the same period, which matches the lower end of this range, and seizures understate the growth of production.

The inversion. The rise in the legal premium tells us how many chemists the legal market lost. The estimated legal premium is $\widehat{d \ln \pi_L} = 0.0589$. In the model, the shock enters only the illegal market: a higher illegal premium draws chemists out of the legal market, legal supply falls, and the legal premium rises until the legal market clears again. For the legal premium to rise by 5.9 log points, the legal market must lose about $(\sigma_L + \alpha_L) \times 5.9 \approx 6.5\%$ of its chemists. The legal-market condition then gives the illegal premium $d \ln \pi_I$, and the illegal-market condition gives the demand shift $d \ln A$ needed to absorb those chemists.[20]

Each implied magnitude depends mainly on one of the two unknown quantities. The implied illegal premium depends on κ_L and does not depend on r, because it is set by how hard it was to draw chemists out of the legal market. The implied demand shift depends on r and very little on κ_L: the number of chemists who moved is set by the legal market, and r converts it into a growth rate of the illegal sector. Figure 10 therefore plots the illegal premium against κ_L in Panel (a), and the demand shift against r in Panel (b).

Parameters. Most parameters come from the paper's estimates or from the literature, and the two without benchmarks vary along the axes of the figure. The supply elasticity of legal chemists α_L follows from the enrollment estimate. Chemistry enrollment responds to the premium with an elasticity of $0.057/0.0589 \approx 0.97$, but only two to five treated cohorts enter the workforce by the end of the period, out of a working life of about 35 cohorts. The elasticity of the stock of chemists is therefore $\alpha_L \approx 0.97 \times (2\text{–}5)/35 \in [0.06, 0.14]$, and I use 0.10. The legal demand elasticity σ_L is set to 1, the upper end of estimated labor demand elasticities, which range from 0.3 to 1. The implied magnitudes increase with σ_L, so this choice is conservative, and the shaded band in the figure covers $\sigma_L \in [0.75, 1.5]$. The illegal-side elasticities ($\sigma_I + \alpha_I = 0.85$) have little effect on the results. The reallocation elasticity κ_L has no estimate in the literature, so Panel (a) covers values

[20] From the legal-market condition, using the identity $-\varepsilon_{LI} = \kappa_L \lambda$, the pass-through ψ gives $d \ln \pi_I = \widehat{d \ln \pi_L}/\psi = \left[(\sigma_L + \alpha_L + \kappa_L)/(\kappa_L \lambda)\right] \widehat{d \ln \pi_L}$, where $\lambda \equiv (1 - p(\bar{\theta}_C))\, \pi_I/\pi_L = 1 - R(\bar{\theta}_C)/\pi_L \in (0, 1]$ (Case A). From the illegal-market condition, $d \ln A$ follows exactly; because $\kappa_I = \kappa_L \lambda / r$ is large when the illegal sector is small, it is well approximated by $d \ln A \approx \left[(\sigma_L + \alpha_L)/r\right] \widehat{d \ln \pi_L}$ plus order-one terms.

from 0.2 to 1.5.[21] The relative size of the illegal sector r covers 1 to 5% in Panel (b). Finally, I set $\lambda = 0.8$; it enters only through $\kappa_L \lambda$, so a different value only moves the results along the κ_L axis of Panel (a).

Implied magnitudes. The illegal chemistry premium must have risen by 14 to 61%, and by 27% at the central value of κ_L (Figure 10, Panel (a)). These increases are within the range suggested by reports of very high pay for chemists recruited by cartels. The higher values correspond to low reallocation, and the rise in employment in legal chemistry occupations (Section 8) suggests that reallocation was not low. Panel (b) shows that the implied growth in illegal demand depends on the size of the illegal sector. If illegal production employed about 2% as many chemists as the legal sector, illegal demand for chemistry skills must have grown about thirtyfold over 2019–2024. At 5%, it must have grown several-fold. A small illegal sector that grows quickly can therefore move the legal premium, because the supply of chemists is inelastic in the short run.

External benchmarks. Independent evidence places the illegal sector at roughly 1 to 5% of legal chemistry employment, which is large enough. The sector does not need to be large, but it cannot be negligible: below about 1%, explaining the legal premium would require an implausibly large shock. Three pieces of evidence suggest it is above that level.

- *Legal chemistry employment.* The average exposed commuting zone employs about 675 workers in chemistry-related occupations, about 92,000 in total.[22] For r to reach 1 to 5%, illegal production needs only 7 to 34 chemists per zone, or 900 to 4,600 in total. Hundreds of clandestine laboratories, each with a few skilled cooks, easily reach this number.

- *Criminal employment.* Prieto-Curiel et al. (2023) estimate that Mexican cartels employed 160,000 to 185,000 people in 2022, about 0.3% of the national workforce, concentrated in production areas. Relative to a criminal workforce of this size, a few thousand chemists is a small share.

- *Growth of production.* Comparing the years after the ban with the years before, and leaving out 2019, laboratory seizures grew about 4.4-fold and fentanyl seizures about 8.7-fold. Since

[21] Reallocation cannot quickly create new chemists able to synthesize drugs, so the low short-run elasticity applies to the skilled workers who earn the premium. Less-educated workers who move into chemistry-related jobs raise employment in chemistry occupations (Section 8) without adding synthesis skills.

[22] Author's calculations from ENOE, 2015–2018, exposed commuting zones, survey-weighted. Chemistry-related occupations comprise laboratory, chemical, pharmaceutical, and agro-industrial occupations and chemistry teaching. The estimate excludes Mexico City, whose commuting zone has about 90,000 such workers; including it raises the average to about 1,325 per zone.

enforcement was weak (Section 8.1), these are lower bounds on the growth of production. A demand shock of this size explains the legal premium for values of r in the low single digits (Figure 10, Panel (b)), in line with the employment evidence.

Internal consistency. The education responses are also of a plausible size relative to the premium. Dividing each schooling response by the rise in the premium gives elasticities of school completion with respect to the chemistry premium of 0.52 at high school age and 0.22 at middle school age, and an elasticity of chemistry-degree enrollment of $0.057/0.0589 \approx 0.97$.[23] All three are positive. They are largest for the choice of field and smallest at middle school age, where almost all children are already enrolled, and they are within the range of schooling responses to returns documented in the literature.

Caveats. The exercise shows that the estimated premium and education responses are consistent with the mechanism under reasonable parameter values. It could have failed: a much larger education response relative to the premium, or an implied illegal premium or demand shock outside any plausible range, would have been hard to reconcile with the model. The calculations assume Case A of the model and an active sorting margin, which the labor-market evidence supports but does not prove.

10 Conclusion

This paper shows that a skill-biased shift in the technology of an illegal market can affect human-capital decisions of the youth exposed to it. China's 2019 fentanyl ban pushed Mexican criminal organizations into domestic synthesis of fentanyl, making clandestine drug production in Mexico more chemistry-intensive and raising the local demand for chemistry skill in the commuting zones with prior synthetic-drug infrastructure. Comparing those exposed zones against corresponding controls in a difference-in-differences design, I find that exposure reduced school dropout, raised educational attainment, and shifted tertiary enrollment specifically toward chemistry-intensive fields, with no change in non-chemistry STEM fields or in overall university enrollment. A model of a local market for chemistry skill organizes these responses and connects them to observable labor-market evidence: the legal wage premium for chemistry-relevant degrees rose and employment in

[23] The completion responses are the reductions in dropout of 2.3 and 1.2 percentage points divided by the control completion rates of 75.2% and 94.3%. They combine the responses of youth who respond through the legal premium, youth who switch sector, and youth who respond through the illegal premium, so the ratio reflects an average over these groups.

chemistry occupations expanded. At the same time, the observed chemistry workforce became less credentialed and the observable local stock of degree-holding chemists shrank, consistent with the most skilled being drawn into illegal production.

The central contribution is that criminal activity does not always discourage human capital accumulation; the skill content of the criminal task matters. The expansion of this criminal activity increased the demand for educated individuals and, crucially, occurred under low enforcement and stable cartel control, so it arrived without a rise in violence. Through the lens of the model, the mechanism is the increased demand for skill from the illegal sector, transmitted to the legal one through a shared labor market and input linkages between illegal and legal operations. Because the short-run supply of chemists is nearly fixed, this transmission does not require a large illegal sector: one employing only a few percent of local chemists is enough to produce the observed rise in the legal premium. The evidence is also suggestive that the illegal sector competes for credentialed workers, so that part of the skill upgrading may be captured by the criminal organizations themselves.

These findings qualify the view of education as a protective factor against crime. When criminal organizations reward the skills that schools teach, education is less effective at making legal alternatives more attractive than criminal ones, and its protective effect may be limited. This lesson extends beyond synthetic drugs: criminal organizations increasingly rely on skills that legal employers also value, such as financial and legal expertise to launder proceeds and information technology and language skills to run online fraud. Where criminal and legal employers compete for the same skills, expanding education and access to good legal jobs also expands the pool of skilled workers that criminal organizations can recruit. Crime prevention in these settings has to be complemented with other tools, such as enforcement against the criminal organizations and the legal firms that supply and support their production, as well as the social and cultural norms that can make criminal work less attractive.

The findings also carry a new lesson for drug policy. Supply-side interdiction is usually evaluated through its effect on prices, quantities, and—increasingly—violence. This evidence points to a further margin: interdiction can push criminal organizations *up the skill ladder*. Before the ban, fentanyl was for the most part synthesized in Chinese chemical firms—a legal, industrial process—and only smuggled in finished form in small shipments; the chemical skill was outside the criminal organizations entirely. Banning the finished drug in China did not eliminate its demand in the U.S.; instead, it relocated synthesis into clandestine Mexican laboratories, moved it back to precursors and pre-precursors, and forced the criminal organizations to internalize the chemical expertise—and the accident and exposure risks—previously borne by legal firms. Interdiction can therefore

leave criminal organizations more technologically sophisticated than it found them, and potentially make these organizations harder to dislodge in the long run.

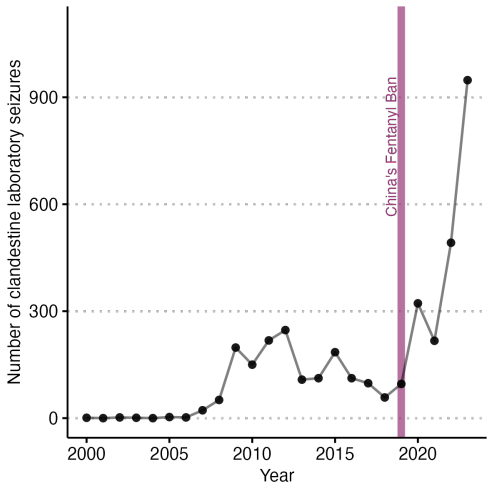
The welfare consequences look different at the individual and the aggregate level. Privately, illegal-bound youth who were always going to enter illicit production plausibly benefit from being pushed into getting more education, insofar as this means moving into a less dangerous, better-paid position, such as switching from working as a soldier to being a lab cook. The young people induced to switch from legal to illegal work take on arrest risk and forgo a legal career, so their private gains are far more ambiguous and plausibly negative once the risk and the criminalization are counted. For the much larger group who simply acquire more schooling in response to the higher legal return, the effect is likely positive—provided the local labor market can absorb the additional chemistry skills rather than leaving them idle or mismatched.

At the aggregate level, the human-capital gains documented here are bound up with the strengthening of the criminal organizations that generated the demand: the same process that keeps youth in school also supplies the illicit sector with a more skilled workforce and a more advanced production capability. Whether the net effect is positive will depend on how one weighs an increase in local human capital against a more entrenched and more sophisticated criminal economy, and on whether the education gains persist in this context.

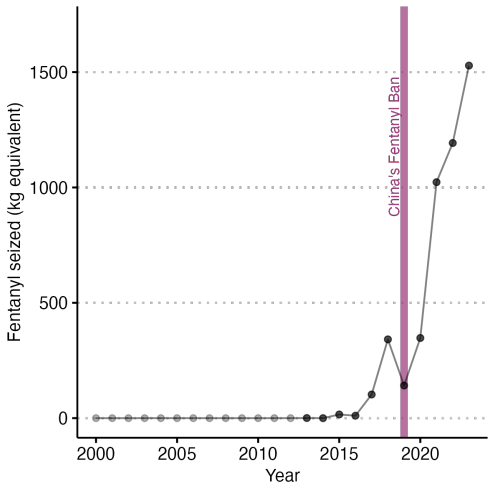
Figures and Tables

Figure 1: Illegal drug seizures in Mexico

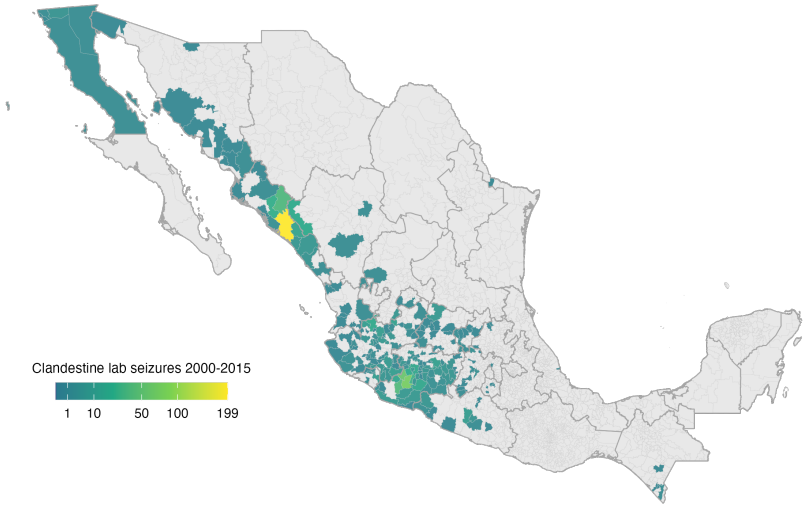
(a) Clandestine laboratory seizures



(b) Fentanyl seizures, kilograms



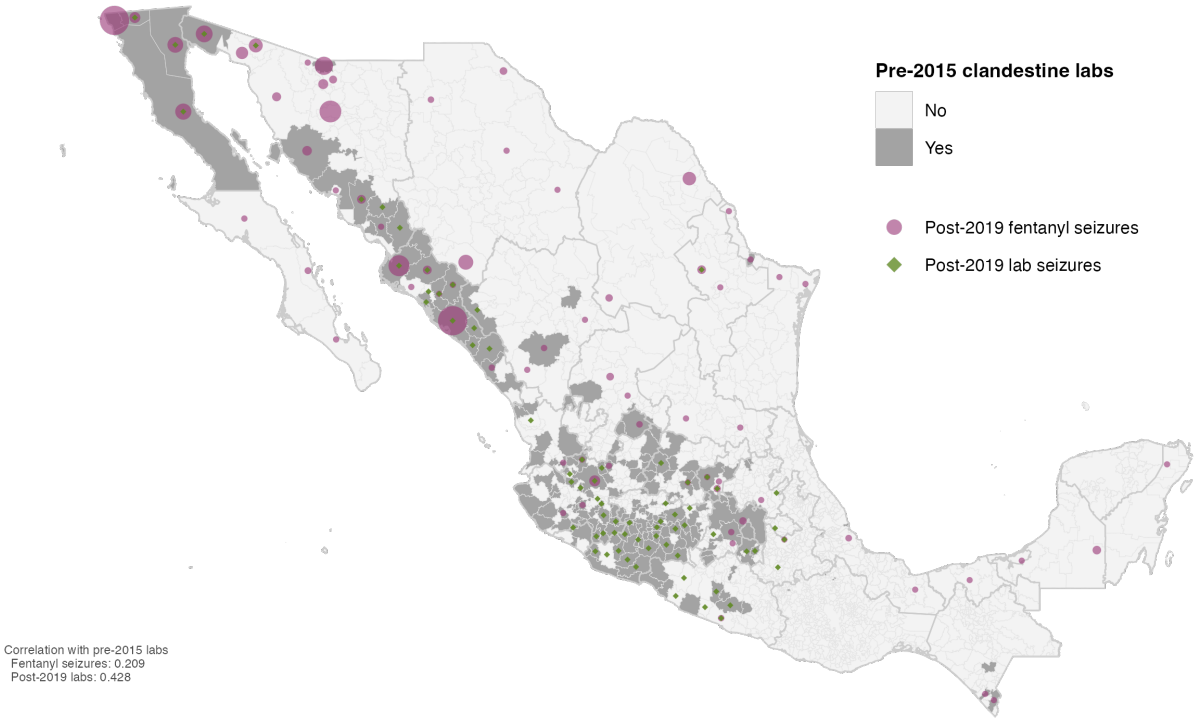
(c) Map of clandestine laboratories seized before 2015, by municipality



Notes: Panels (a) and (b) show the number of clandestine laboratory and kilogram amount of fentanyl seizures by year, respectively. Vertical shaded areas indicate Mexican presidential administrations: VFQ (Vicente Fox Quesada, 2000-2006), FCH (Felipe Calderón Hinojosa, 2006-2012), EPN (Enrique Peña Nieto, 2012-2018), and AMLO (Andrés Manuel López Obrador, 2018-2024). The vertical line marks 2019, when synthetic drug production became widespread. Panel (c) shows the spatial distribution of clandestine laboratory seizures from 2000-2015 at the municipality level.
Source: Author's calculations based on Mexican Secretariat of National Defense (SEDENA) seizure records obtained via Insight Crime.

⟷ Return to text

Figure 2: Spatial distribution of treatment and post-treatment drug production

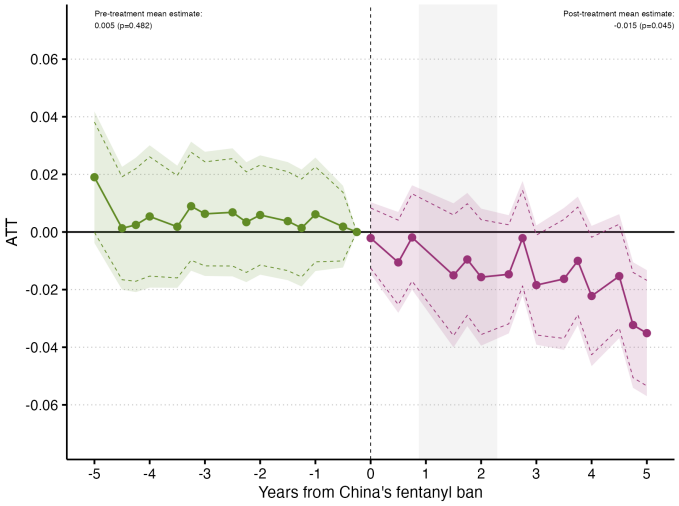


Notes: The map shows the geographic distribution of the main treatment variable and subsequent drug production activity. Gray shading indicates treatment status: commuting zones with high exposure (dark gray) had clandestine laboratory seizures before 2015, while low exposure zones (light gray) had no pre-2015 labs. Purple circles show fentanyl seizures occurring after the Chinese fentanyl ban, with circle size proportional to kilograms seized. Green diamonds indicate post-2019 clandestine laboratory seizures. Correlations demonstrate that drug production activity persists in treated regions: areas with pre-2015 exposure experience higher rates of both fentanyl seizures and new laboratory discoveries. *Source:* Author's calculations based on Mexican Secretariat of National Defense (SEDENA) press releases and seizure records.

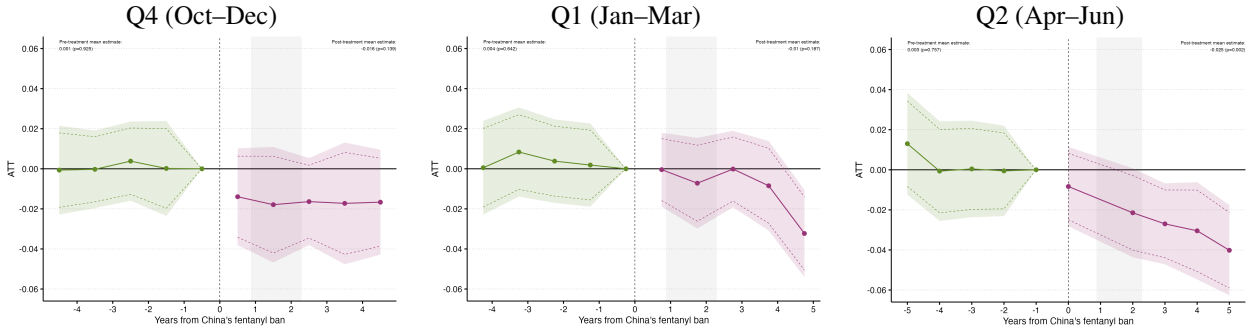
↩ Return to text

Figure 3: Event Study of Exposure to Synthetic Drug Production on High School Dropout

(a) All quarters



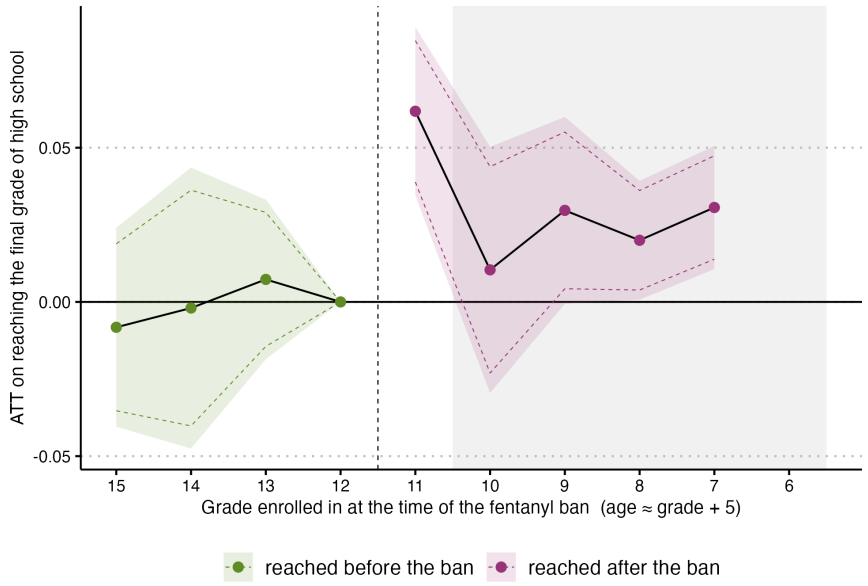
(b) By quarter



Notes: Event-study estimates of the effect of exposure on school dropout at high school age (15–17). The dependent variable is an indicator for not being enrolled in school. Exposure is an indicator for any methamphetamine laboratory seized in the commuting zone before 2015. Panel (a) pools all quarters: exposure is interacted with quarter indicators, omitting the quarter immediately before the fentanyl ban, and the specification includes commuting zone, year-quarter, and quarter fixed effects. Panel (b) estimates the event study separately for each calendar quarter of the survey: Q4 (Oct–Dec, start of the Mexican academic year), Q1 (Jan–Mar), and Q2 (Apr–Jun, end of the year), with exposure interacted with year indicators relative to the last pre-ban observation of each quarter and commuting zone and year fixed effects. All specifications control for gender, age, and household education, and observations are weighted by survey weights. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level. The shaded region marks the COVID-19 school-closure period. Middle school results are in Appendix G.1.
Source: Author's calculations based on ENOE.

↩ Return to text

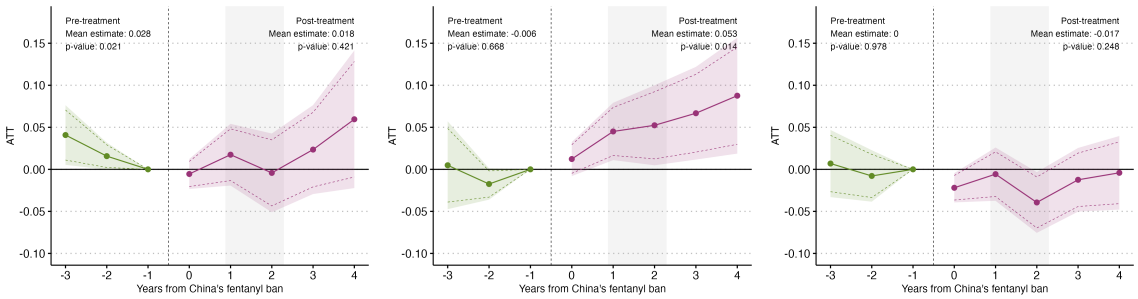
Figure 4: Effect of Exposure to Synthetic Drug Production on Reaching the Final Grade of High School



Notes: The figure plots coefficients from a cohort event study estimated on the Formato 911 school census. The dependent variable is the share of a middle-school entry cohort in a commuting zone that remains continuously enrolled through the final grade of high school (a six-grade trajectory). Cohorts are indexed by the grade they were enrolled in when the fentanyl ban was announced in April 2019 (x-axis; age $\approx$ grade $+ 5$), and the specification includes commuting-zone and grade-at-ban fixed effects, weights cohorts by size, and clusters standard errors at the commuting-zone level. The omitted category is the cohort enrolled in the final grade at the ban (the last to reach it before the shock). Green markers denote cohorts that reached the final grade before the ban and pink markers those that reached it in 2019 or later; the vertical dashed line marks the ban, and the shaded region marks cohorts that reached the final grade during or after the COVID-19 school closures. The single unshaded post-ban cohort (grade 11 at the ban) reached the final grade in September 2019—after the ban but before the closures. Bands are 95% confidence intervals and dashed lines 90% confidence intervals.
Source: Author's calculations based on Formato 911.

↩ Return to text

Figure 5: Impact of the Synthetic Drug Expansion on Tertiary enrollment

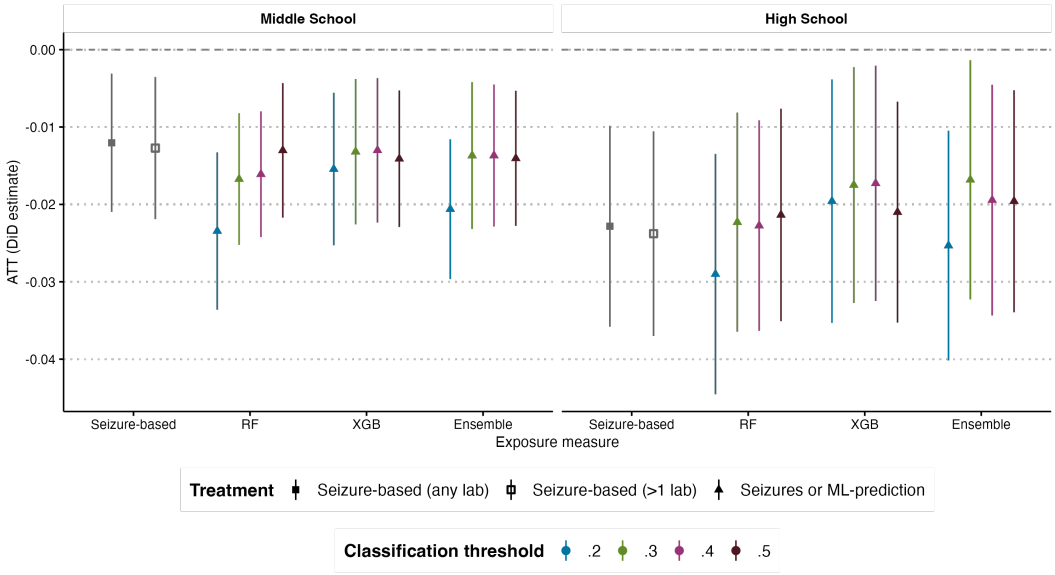


Notes: Event-study estimates of the effect of exposure on tertiary (university) enrollment by field of study, at the commuting-zone level. Panels show total enrollment, chemistry-intensive STEM+H programs, and STEM+H programs without chemistry. Estimates are from a Poisson (PPML) regression of field-specific enrollment counts on commuting-zone exposure interacted with year indicators (reference: 2018, the year before the ban), with commuting zone and year fixed effects; coefficients are proportional effects (log points). Exposure is an indicator for any methamphetamine laboratory seized in the commuting zone before 2015. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level.
Source: Author's calculations based on ANUIES tertiary enrollment data.

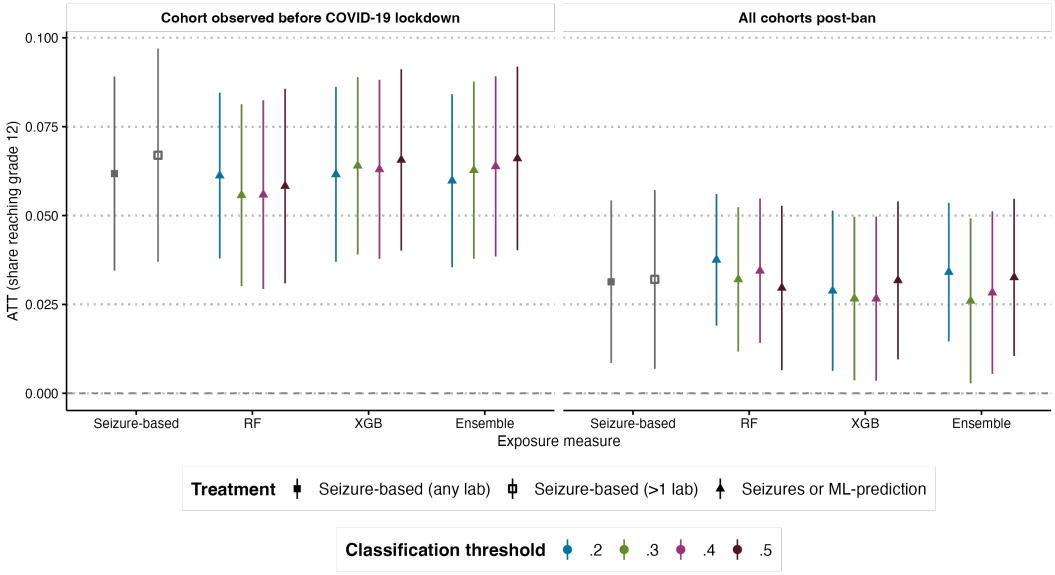
↩ Return to text

Figure 6: DiD estimates of alternative treatment measures

Panel A. School dropout



Panel B. Share of cohort reaching grade 12



Panel C. University enrollment

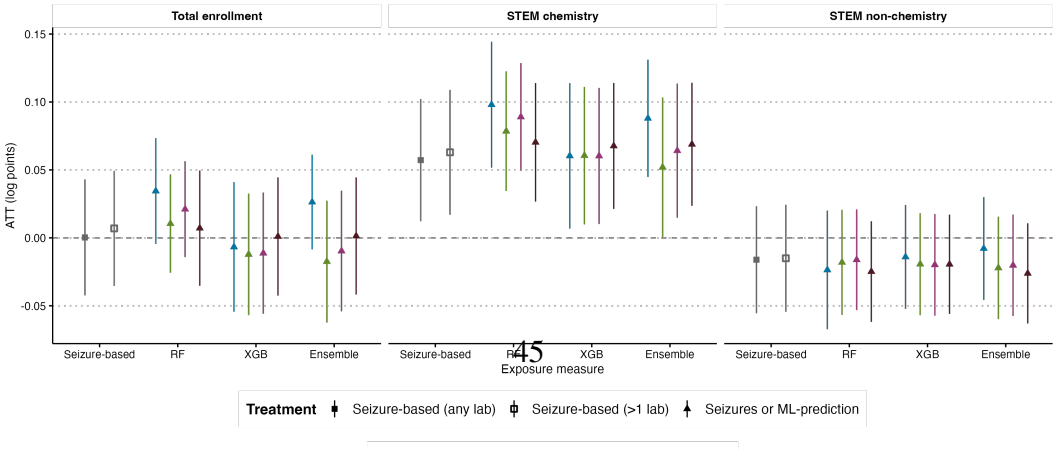
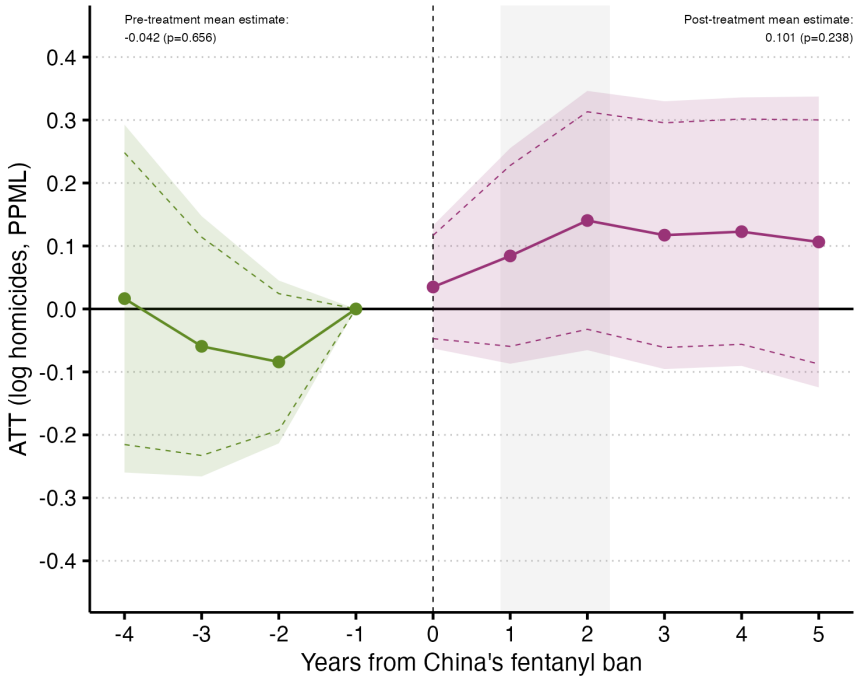


Figure 7: Exposure and the intentional-homicide rate

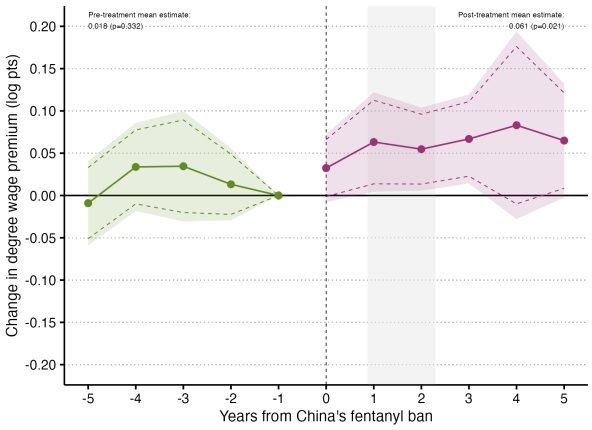


Notes: Event-study estimates of exposure on intentional homicides at the commuting-zone × year level, 2015–2024, relative to the year before the ban ($t = -1$). Estimated by Poisson pseudo-maximum-likelihood with a log-population offset, so coefficients are proportional changes in the per-capita homicide rate; population is the annual commuting-zone total from CONAPO municipal projections. Specification includes commuting-zone and year fixed effects; standard errors clustered by commuting zone. Robustness to alternative population measures and to the high-violence subsample is reported in Appendix Table J.1 and Appendix Figure J.1.

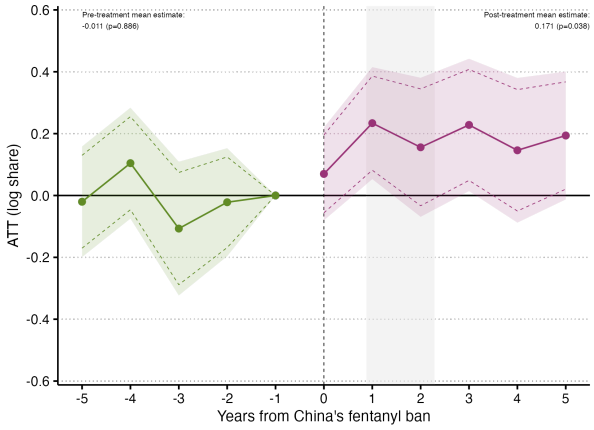
↩ Return to text

Figure 8: Event studies: the price and quantity of chemistry skill

(a) Return to a STEM degree (price)



(b) Chemistry-occupation employment (quantity)

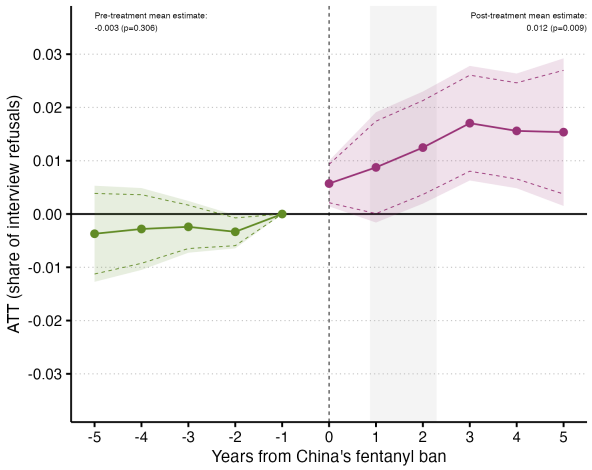


Notes: Event studies of the two headline labor-market outcomes, relative to the year before the ban ($t = -1$). Panel (a) is the change in the wage premium on a STEM degree, from a triple-difference on log hourly wages; the chemistry-specific estimate is in Table 4, and the broader STEM series is shown here as the cleaner parallel-trends check. Panel (b) is the share of employment in chemistry occupations, estimated by Poisson pseudo-maximum-likelihood at the commuting-zone × year level with a log-employment offset. Both specifications include commuting-zone and year (survey-wave) fixed effects; standard errors clustered by commuting zone. Green markers are pre-treatment estimates, purple post-treatment; shaded bands are 95% confidence intervals; the shaded vertical region marks the COVID-19 school-closure window.

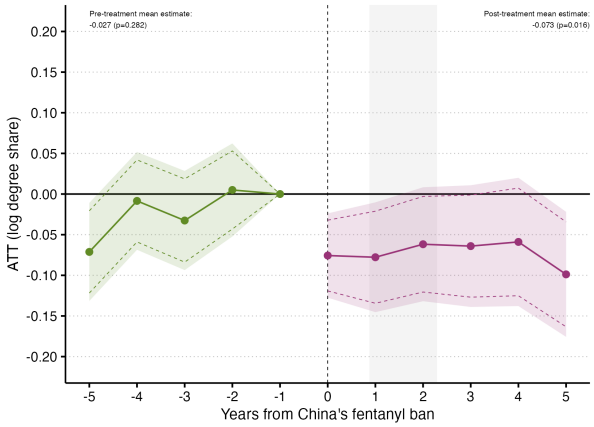
↩ Return to text

Figure 9: Exposure, survey refusals, and the observed stock of STEM degree holders

(a) Household refusal rate



(b) STEM-degree share of the fixed 2019 cohort



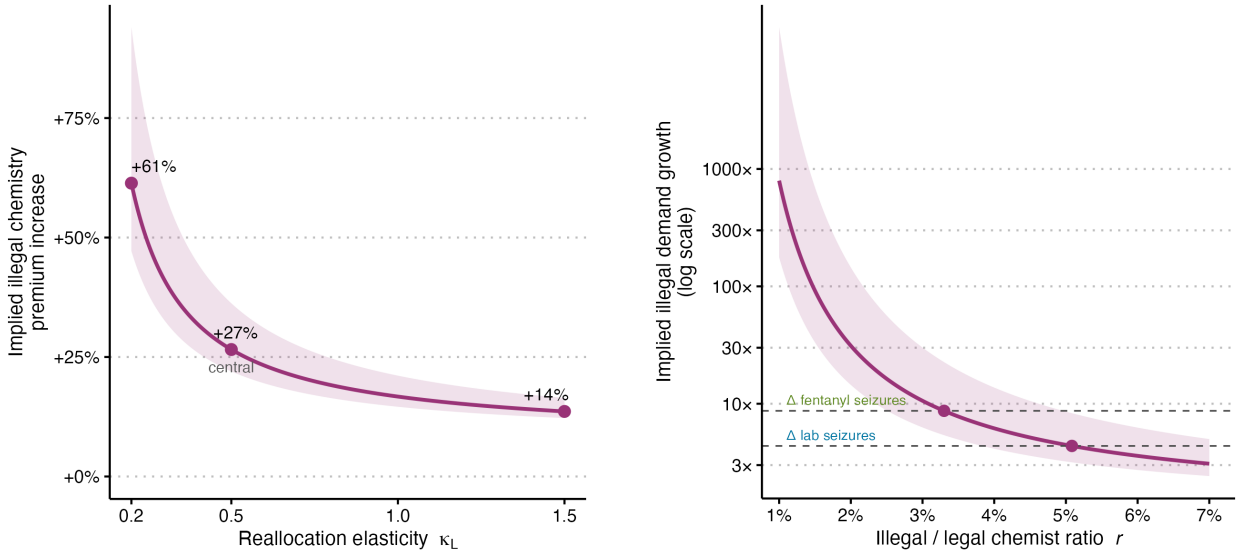
Notes: Event studies relative to the year before the ban ($t = -1$). Panel (a) is the household refusal rate—the share of inhabited dwellings whose occupants decline the ENOE interview—at the commuting-zone × year level. Because ENOE expansion factors are zero for non-respondents, the panel is unweighted; the specification includes commuting-zone and year fixed effects, a seasonal (quarter) fixed effect, and commuting-zone linear trends. Panel (b) is the share of the fixed 2019 cohort (aged 30–54 in 2019) holding a STEM degree, estimated by Poisson pseudo-maximum-likelihood at the commuting-zone × year level with an offset for log cohort size and commuting-zone and year fixed effects. Standard errors clustered by commuting zone. Green markers are pre-treatment estimates, purple post-treatment; shaded bands are 95% confidence intervals; the shaded vertical region marks the COVID-19 school-closure window.

↩ Return to text

Figure 10: Model-implied illegal-sector magnitudes from the estimated legal chemistry premium ($\widehat{d \ln \pi_L} = 0.0589$)

(a) Implied illegal chemistry premium, by reallocation elasticity κ_L

(b) Implied illegal demand growth, by relative sector size r



Notes: Both panels invert the difference-in-differences estimate of the legal chemistry premium ($\widehat{d \ln \pi_L} = 0.0589$) through the market-clearing conditions of Section 3. Panel (a) plots the implied rise in the *illegal* chemistry premium against the reallocation elasticity κ_L; Panel (b) plots the implied growth in illegal demand for chemistry skill against the illegal-to-legal chemist ratio r (log scale). The solid line holds the remaining parameters at central values ($\sigma_L = 1$, $\alpha_L = 0.10$, $\lambda = 0.8$, $\sigma_I + \alpha_I = 0.85$); the shaded band traces the legal demand elasticity over $\sigma_L \in [0.75, 1.5]$ and is a parameter-robustness range, not a confidence interval. In Panel (b), the dashed lines mark the growth of detected clandestine-laboratory (Δ lab) and fentanyl (Δ fentanyl) seizures over the study window (average 2020–2023 relative to average 2015–2018, dropping the 2019 transition year), and the solid points their implied r; because seizures undercount production, these growth factors are lower bounds and the implied r upper bounds.
Source: Author's calculations.

↩ Return to text

Table 1: Validating Exposure Measures: Future Drug Lab Seizures and Drug Production Byproducts in Water

	Dependent variable:					
	Lab seizures post-2015		Mercury detection		Toluene detection	
	RF	XGB	RF	XGB	RF	XGB
Panel A. Classification threshold $\tau = 0.20$						
Observed exposure	0.095***	0.096***	0.116***	0.103***	0.024***	0.020***
	(0.013)	(0.013)	(0.016)	(0.016)	(0.005)	(0.005)
Predicted exposure	0.030***	0.030***	0.051***	0.076***	0.015***	0.022***
	(0.005)	(0.006)	(0.011)	(0.011)	(0.004)	(0.004)
Panel B. Classification threshold $\tau = 0.50$						
Observed exposure	0.088***	0.089***	0.148***	0.117***	0.032***	0.025***
	(0.013)	(0.013)	(0.014)	(0.016)	(0.005)	(0.005)
Predicted exposure	0.170***	0.050***	-0.006	0.059***	0.005	0.015***
	(0.047)	(0.011)	(0.022)	(0.014)	(0.010)	(0.005)
Mean dep. var.	0.015	0.015	0.312	0.312	0.073	0.073
Observations	17,700	17,700	10,111	10,111	8,326	8,326

Notes: Each column reports estimates from an OLS regression at the commuting zone (CZ)–year level. All specifications include year fixed effects and cluster standard errors at the CZ level. "Observed exposure" equals one if any synthetic drug laboratory was seized in the CZ before 2015. "Predicted exposure" equals one if the random forest (RF) or gradient-boosting (XGB) model, respectively, assigns an out-of-fold predicted probability of laboratory presence above the classification threshold τ: $\tau = 0.20$ in Panel A and $\tau = 0.50$ in Panel B. Columns 1–2 take as outcome an indicator for any laboratory seizure in CZ c in year t, over 2015–2024, unweighted. Columns 3–6 use water quality outcomes over 2015–2020: the share of monitoring sites in the CZ where mercury or toluene concentrations exceeded detection limits, aggregated with inverse-distance site-to-CZ contiguity weights and weighted by them in estimation. These four columns additionally control for the standardized chemical-to-biological oxygen demand ratio (COD/BOD), which measures general industrial pollution, and for the number of monitoring sites in the CZ. The toluene columns have fewer observations because toluene is tested at fewer sites. Mean dep. var. and the number of observations are common to both panels. At $\tau = 0.50$ the random forest (gradient-boosting) classifies 34 (179) of 1,771 CZs as exposed, against 335 (283) at $\tau = 0.20$. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

↩ Return to text

Table 2: Effect of Exposure to Synthetic Drug Production on High School Dropout

	All quarters				By quarter		
	(1)	(2)	(3)	(4)	Oct-Dec	Jan-Mar	Apr-Jun
Exposure × Post	-0.021**	-0.022**	-0.023***	-0.021**	-0.020*	-0.017**	-0.031***
	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)	(0.006)	(0.007)
Exposure × Post × Female				-0.003			
				(0.009)			
Female		-0.028***	-0.029***	-0.012***	-0.024***	-0.029***	-0.034***
		(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
Age		0.054***	0.053***	0.053***	0.055***	0.053***	0.050***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Household avg. education			-0.010***	-0.010***	-0.009***	-0.010***	-0.010***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	674,437	674,437	674,437	674,437	218,077	234,183	222,177
Mean of dep. var.	0.248	0.248	0.248	0.248	0.230	0.249	0.264

Notes: The dependent variable is an indicator for not being enrolled in school among individuals of high school age (15–17). Exposure is an indicator for any methamphetamine laboratory seized in the commuting zone before 2015. Post is an indicator for periods from the second quarter of 2019. Household avg. education is the average years of schooling of adults older than 25 in the individual's household. Columns (1)–(4) pool all quarters and include commuting zone, year-quarter, and quarter fixed effects. Columns (5)–(7) are estimated separately for each calendar quarter: Q4 (October–December) corresponds to the start of the Mexican academic year, which runs August–July, and Q2 (April–June) to its end; these columns include commuting zone and year fixed effects and control for gender, age, and household education. Observations are weighted by survey weights. Mean of dep. var. is the mean in control commuting zones before the ban. Middle school results are in Appendix Table G.1. Standard errors clustered at the commuting zone level are reported in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

↩ Return to text

Table 3: Impact of the Synthetic Drug Expansion on tertiary enrollment

	Enrollment			
		STEM+H		
	All fields	Chemistry	Non-Chemistry	Non-STEM+H
Panel A. All				
Exposure × Post	0.000	0.057*	-0.016	-0.028
	(0.022)	(0.023)	(0.020)	(0.026)
N	8,880	8,512	8,128	8,840
Control mean	2644.700	783.900	428.800	1510.300
Panel B. Female				
Exposure × Post	-0.009	0.061**	-0.049+	-0.029
	(0.020)	(0.022)	(0.027)	(0.026)
N	8,880	8,512	8,128	8,840
Control mean	1356.600	353.800	195.500	843.500
Panel C. Male				
Exposure × Post	0.010	0.048+	0.010	-0.024
	(0.024)	(0.025)	(0.019)	(0.027)
N	8,880	8,512	8,128	8,840
Control mean	1288.100	430.000	233.300	666.800

Notes: All models include commuting zone and year fixed effects. Standard errors clustered at the commuting zone level. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

↩ Return to text

Table 4: Effects of synthetic drug production on returns to education and employment

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Effects on wages and returns to education degrees						
	Dependent variable: Hourly wage					
		Degree group interaction				
					STEM by chemistry content	
	Overall wage	Any tertiary	Non-STEM	STEM	No	Yes
Exposure × Post	-0.0121	-0.0211	-0.0195	-0.0135	-0.0129	-0.0127
	(0.0143)	(0.0206)	(0.0187)	(0.0152)	(0.0147)	(0.0148)
Exposure × Post × Degree$_g$		0.0407**	0.0318	0.0501**	0.0234	0.0589**
		(0.0197)	(0.0195)	(0.0222)	(0.0229)	(0.0255)
Observations (ind. × year)	3,655,535	3,655,535	3,655,535	3,655,535	3,655,535	3,655,535
Mean (control)	3.388	3.388	3.388	3.388	3.388	3.388
Baseline premium (log pts)		0.733	0.642	0.514	0.468	0.520
No. of commuting zones	1,230	1,230	1,230	1,230	1,230	1,230
Panel B. Effects on employment and share of employment by occupation						
	Dependent variable: employment (PPML)					
		By occupation group				
					STEM by chemistry content	
	All	Professional	Non-STEM	STEM	No	Yes
Exposure × Post	-0.0015	0.0130	0.0041	0.0733***	0.0514	0.1762***
	(0.0034)	(0.0096)	(0.0107)	(0.0283)	(0.0326)	(0.0606)
Observations (CZ × year)	8,955	8,768	8,756	6,771	6,484	4,434
Mean (control)	0.7568	0.1931	0.1720	0.0215	0.0179	0.0036
No. of commuting zones	1,230	1,152	1,149	729	682	435

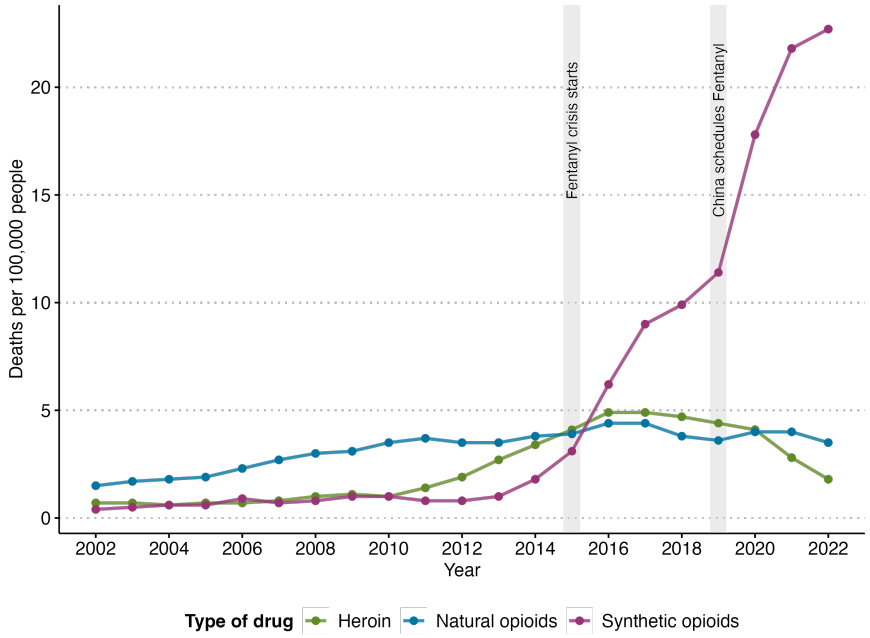
Notes: Difference-in-differences estimates; standard errors clustered by commuting zone in parentheses. *Exposure* equals one for commuting zones with a methamphetamine laboratory seized before 2015; *Post* equals one from 2019 (China's fentanyl ban). Sample: prime-age (25–54). *Panel A* reports OLS estimates on the log hourly wage: column (1) is the overall wage effect, and columns (2)–(6) add a degree-group (g) interaction whose triple interaction ($Exposure \times Post \times Degree_g$) is the change in that degree's wage premium in treated commuting zones after the shock; all columns control for sex and a quadratic in age, include commuting-zone and survey-wave (year × quarter) fixed effects, and are weighted by survey weights. *Panel B* reports Poisson (PPML) estimates at the commuting-zone × year level: column (1) is the employment rate (log-population offset) and columns (2)–(6) are the share of employment in each occupation group (log-employment offset), so coefficients are proportional changes; all include commuting-zone and year fixed effects. The number of commuting zones falls across the columns of Panel B because the Poisson estimator drops commuting zones with no within-zone variation in the outcome—a zone with no employment in a given occupation group in any year is perfectly predicted by its fixed effect and contributes nothing to the likelihood. In both panels columns (5)–(6) split STEM by chemistry content; column (6) is chemistry: chemistry-course fields in Panel A (including pharmacy) and chemistry occupations in Panel B (chemists, chemical engineers, and chemical-plant operators). $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$.

⟵ Return to text

Appendix

A Additional Context

Figure A.1: Evolution of overdoses in the United States by type of drug



Notes: The figure shows overdose death rates per 100,000 people by drug type from 2002-2023. Synthetic opioids include fentanyl and its analogs. Natural opioids include morphine, codeine, and semi-synthetic opioids such as oxycodone and hydrocodone. The first shaded region indicates the introduction of fentanyl into U.S. supplies. The second shaded region marks China's fentanyl ban in 2019, which coincided with accelerated growth in synthetic opioid deaths.
Source: National Institute on Drug Abuse (NIDA), based on CDC Wonder, Multiple Cause of Death data.

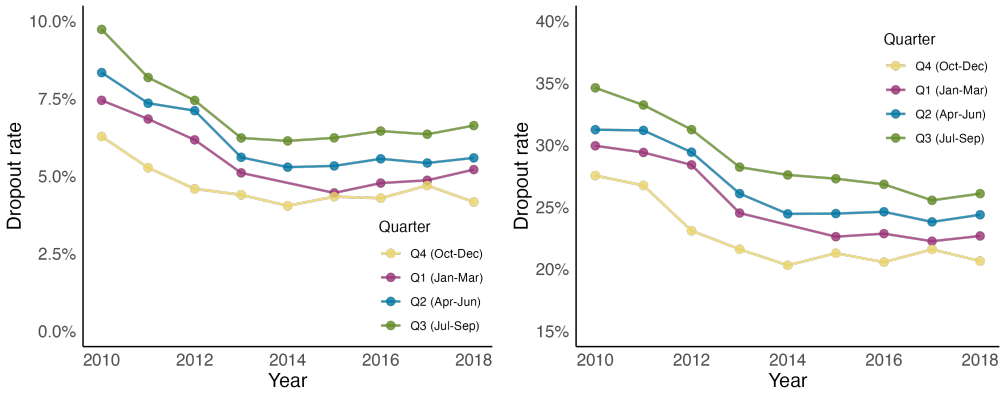
↩ Return to text

B Data

Figure B.1: Seasonality of school dropout

(a) Middle school age

(b) High school age

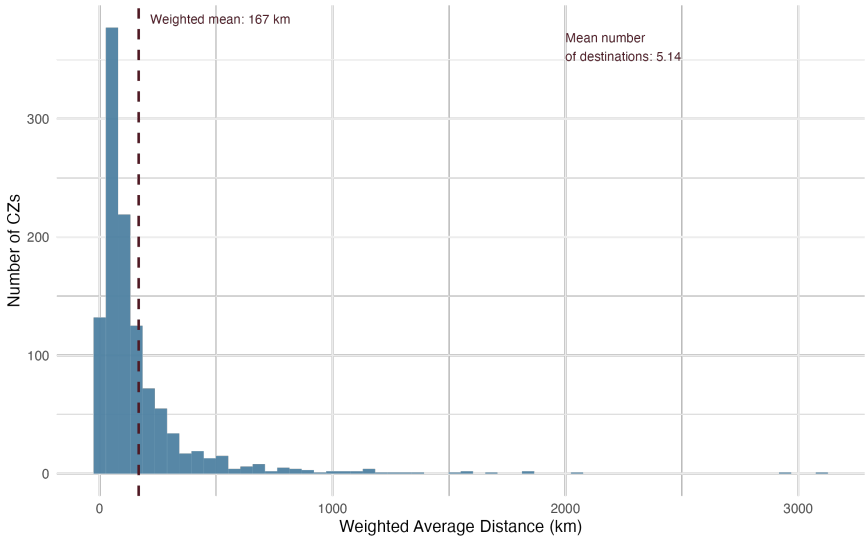


Notes: The figure shows dropout rates by quarter and school year for middle school (ages 12-14, panel a) and high school (ages 15-17, panel b) students in Mexico. Dropout is defined as not currently attending school for children in the relevant school age. The Mexican school year runs from August/September to June/July, with breaks in December-January (Q4/Q1 boundary) and during summer (Q2/Q3 boundary). Q3 (July-September) shows consistently higher dropout rates as students fail to attend school during the summer break, while Q4 (Oct-Dec) shows lower dropout as it reflects the first full quarter of a given academic year.
Source: Author's calculations based on ENOE (National Survey of Occupation and Employment) quarterly household data, 2010-2018.

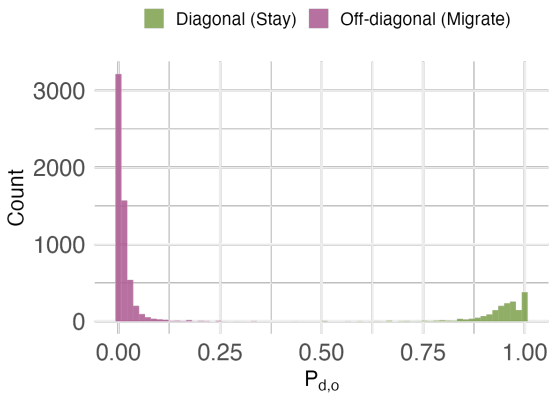
↩ Return to text

Figure B.2: Student migration patterns across commuting zones

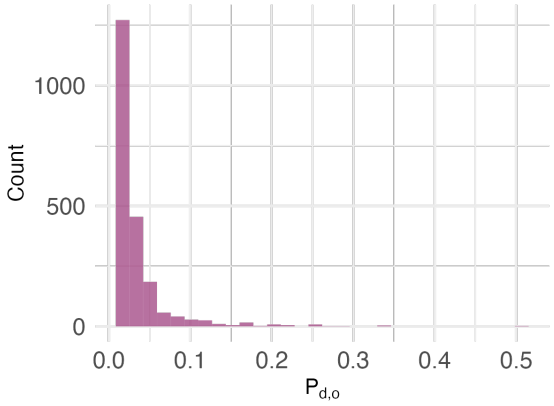
(a) Average migration distance by origin commuting zone, weighted by destination probability



(b) Distribution of probabilities across all origin-destination pairs



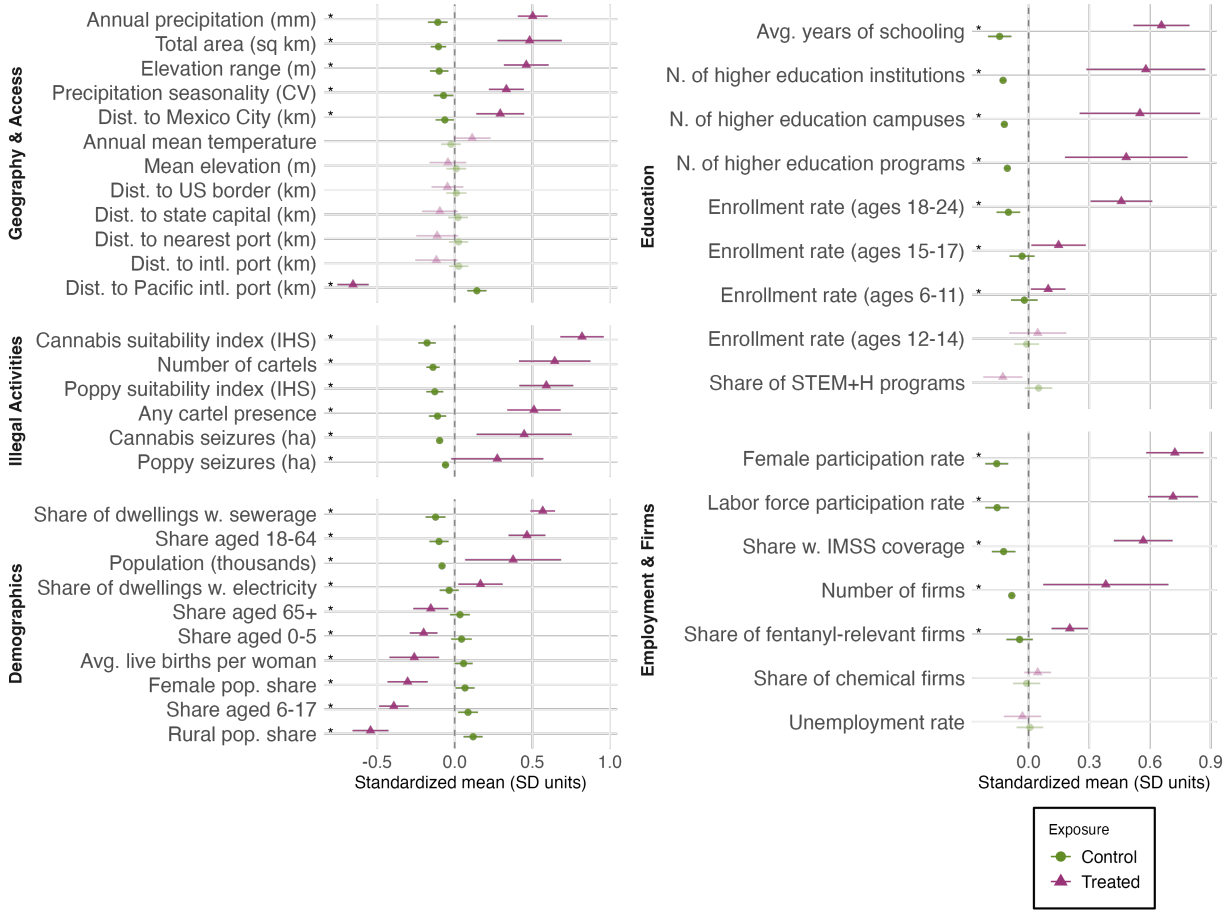
(c) Distribution of migration probabilities, different origin-destination and excluding nonzeros



Notes: The figure characterizes migration patterns between 2005 and 2010 for students aged 19-24. Panel (a) shows the distribution of weighted average migration distances across origin commuting zones. For each origin CZ, the weighted average distance is calculated as the probability-weighted mean distance to all destination CZs (excluding staying in the origin). Distance is measured in kilometers using the Haversine formula between CZ centroids. Panel (b) shows the distribution of migration probabilities $P_{d,o}$ across all origin-destination pairs, distinguishing between diagonal elements (staying in origin CZ, shown in green) and off-diagonal elements (migrating to different CZ, shown in purple). Panel (c) shows the distribution of migration probabilities for off-diagonal pairs only, excluding pairs with near-zero probability (prob > 0.01).
Source: Author's calculations based on migration flows derived from Mexican Population Census 2010 long-form data.

↩ Return to text

Figure B.3: Balance on covariates across treatment and control CZ



Notes: Each point reports the mean of the indicated baseline (pre-2019) commuting-zone characteristic, standardized to mean zero and unit standard deviation, shown separately for treated commuting zones (any methamphetamine laboratory seized before 2015) and control zones; horizontal bars are 95% confidence intervals. Asterisks denote characteristics for which the treated–control difference is statistically significant in a two-sample t-test. Characteristics are grouped into geography and access, illegal activities, demographics, education, and employment and firms.

↩ Return to text

Table B.1: Effect of Exposure on Reaching the Final Grade of High School, by Cohort

Grade enrolled in at the ban	Age at the ban	Reaches final grade in	ATT
Reached final grade before the ban			
15	20	2015	−0.008
			(0.016)
14	19	2016	−0.002
			(0.023)
13	18	2017	0.007
			(0.013)
12	17	2018	— (ref.)
Reached final grade after the ban			
11	16	2019	0.062***
			(0.014)
10	15	2020	0.010
			(0.020)
9	14	2021	0.030+
			(0.015)
8	13	2022	0.020*
			(0.010)
7	12	2023	0.031**
			(0.010)
Pooled post-ban ATT			0.031**
			(0.012)
Pre-shock control mean			0.549
Observations (cohort × CZ)			10,917
Commuting zones (treated)			1,218 (163)

Notes: Coefficients from a cohort event study estimated on the Formato 911 school census. The dependent variable is the share of a middle-school entry cohort in a commuting zone that remains continuously enrolled through the final grade of high school. Each cohort is indexed by the grade its students were enrolled in when the fentanyl ban was announced (April 2019); the omitted category is the cohort enrolled in the final grade (grade 12) at the ban. *Exposure* equals one for commuting zones with a methamphetamine laboratory seized before 2015. The pooled post-ban ATT replaces the grade-at-ban interactions with a single *Exposure* × *Post* term, where *Post* indicates cohorts reaching the final grade in 2019 or later. All specifications include commuting-zone and grade-at-ban fixed effects, weight cohorts by size, and cluster standard errors at the commuting-zone level. Cohorts reaching the final grade in 2020 or later did so during or after the COVID-19 school closures; the grade-11 (2019) cohort reached it after the ban but before the closures. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

↩ Return to text

C Model

This appendix develops in full the model summarized in Section 3: the decision rules, the market for chemistry skill, existence and uniqueness of equilibrium, the comparative statics behind the four predictions, and the analytical magnitude results. The numerical back-of-the-envelope inversion is carried out separately in Section 9. Table C.1 summarizes the notation.

Table C.1: Elements of the model

Symbol	Meaning
Choices (per individual)	
$e \in \{0, 1\}$	complete an additional level of schooling
$f \in \{C, N\}$	field of study: chemistry-intensive (C) or not (N)
$s \in \{L, I\}$	sector of employment: legal (L) or illegal (I)
Individual heterogeneity	
θ	type; higher θ lowers arrest risk (connections, know-how)
$p(\theta)$	arrest probability, $p'(\theta) < 0$
η_i	idiosyncratic net cost of studying the chemistry field
Prices (equilibrium objects)	
$\pi_L,\ \pi_I$	chemistry premium in the legal / illegal sector
Parameters	
$w_{s0},\ \beta_s^N$	base wage and non-chemistry return in sector s
$P,\ c$	penalty on arrest; cost of schooling
$A,\ B$	illegal / legal demand shifter for chemistry skill
$\phi = d \ln B / d \ln A$	input-linkage pass-through ($\phi = 0$: pure poaching)
$\sigma_L,\ \sigma_I$	legal / illegal demand elasticity
κ_L	reallocation elasticity of legal chemistry supply
α_L	creation elasticity (stock of legal chemists)
$r = H_I^s / H_L^s$	size of the illegal relative to the legal chemistry workforce

Notes: The shock studied in the paper—China's 2019 fentanyl ban—is an outward shift in illegal demand for chemistry skill, an increase in A in exposed commuting zones. The two chemistry premiums π_L, π_I are determined by clearing a local market for chemistry-skilled labor in which both sectors bid; $\pi_I > \pi_L$ in equilibrium because illegal work carries arrest risk.

I develop a model of educational choice in which chemistry skill trades in a local labor market served by both legal and illegal employers. The model has two blocks. The first is an occupational-choice block: youth decide whether to stay in school, which field to study, and which sector to work in, given sector-specific premiums for chemistry skill. The second is a market-clearing block: the two premiums — π_L in the legal sector and π_I in the illegal sector — are determined jointly by the supply of chemists that the schooling choices themselves generate and by the two sectors' demands. The skill-biased shock at the studied in the paper, the post-2019 expansion of synthetic drug production, enters as an outward shift in illegal demand for chemistry skill.

The model explains the resulting rise in the *legal-sector* chemistry premium stemming from competition between sectors for a common pool of chemists, and the education response of youth with no propensity for illegal work then follows mechanically, because those youth also face an increased legal premium. The model also disciplines magnitudes: it shows that whether a small illicit sector can move education outcomes is a question about the elasticity of local chemistry-skill supply, not about the sector's headcount, and it yields closed-form expressions that map the reduced-form wage estimates of Section 6 into implied illegal-sector quantities.

Throughout, agents hold *rational expectations*: each individual correctly anticipates the equilibrium premiums (π_L, π_I) and his or her own eventual sector of employment, and takes prices as given. The model is solved by *backward induction*: sector choice first, then field choice, then the schooling decision.

Remark 1 (Notation)**.** *The choice-theoretic notation follows the conventions used throughout the paper: $e \in \{0,1\}$ is the schooling decision, $f \in \{C, N\}$ the field (chemistry-intensive versus other), $\eta_i \sim G$ the idiosyncratic net cost of studying C, θ the type indexing propensity for illegal work, p the arrest probability, P the penalty, c the cost of schooling, and w_{s0}, β_s^N the base wages and non-chemistry returns in sector $s \in \{L, I\}$. New symbols introduced by the market-clearing structure are: the sector-specific chemistry premiums π_L and π_I (replacing β_L^C and β_I^C); the type distribution F with density f_θ, and the density g of G; the cell values $V_0(\theta)$, $V_N(\theta)$, $V_C(\theta; \pi_L, \pi_I)$; the sector thresholds $\bar{\theta}_0$, $\bar{\theta}_N$, $\bar{\theta}_C(\pi_L, \pi_I)$ and the education threshold $\theta_e(\eta_i, \pi_I)$; the crime rent $R(\theta)$ and the chemistry cutoff $\bar{\eta}(\theta; \pi_L, \pi_I)$; sector supplies and demands H_s^{sup}, H_s^d with scale parameters D_L, D_I, demand elasticities σ_L, σ_I, and shifters A (illegal) and B (legal, with pass-through $\phi \equiv d \ln B / d \ln A \geq 0$); the supply-response primitives E_L, E_I (entry margins), m (mass at the sector margin), μ_L, μ_I (sector-margin sensitivities), and δ_L, δ_I (demand slopes); and the supply elasticities $\varepsilon_{LL}, \varepsilon_{LI}, \varepsilon_{IL}, \varepsilon_{II}$ defined in (26). The single supply elasticity ε of a one-price model is the special case in which the sector margin is shut ($m = 0$) and the two premiums coincide.*

C.1 Setup

Consider a continuum of youth in a local labor market (empirically, a commuting zone). Each person is characterized by a pair (θ, η_i) drawn independently their corresponding distrubutions F and G.

Choices and timing. Each individual makes three choices, solved by backward induction: (i) whether to complete an additional level of schooling, $e \in \{0,1\}$, at cost c (reflecting direct plus

opportunity costs of schooling); (ii) if $e = 1$, a field $f \in \{C, N\}$, where C denotes chemistry-intensive programs, at idiosyncratic additional net cost η_i for C; and (iii) a sector of employment $s \in \{L, I\}$.

Wages. Earnings are

$$w_L(e, f) = w_{L0} + e \cdot \begin{cases} \pi_L & f = C \\ \beta_L^N & f = N \end{cases}, \qquad w_I(e, f) = w_{I0} + e \cdot \begin{cases} \pi_I & f = C \\ \beta_I^N & f = N \end{cases}. \tag{9}$$

The non-chemistry returns β_L^N and β_I^N are considered exogenous parameters for simplicity. The chemistry premiums π_L and π_I are *prices*, determined in Section C.3 by clearing in the local market for chemistry skill. Allowing the two premiums to differ is essential: illegal employment carries arrest risk, so in equilibrium the illegal sector must pay a compensating differential to attract chemists, and the size of that wedge is an equilibrium object (Corollary 1).

Arrest risk and types. Illegal work leads to arrest with probability p, in which case the wage is forfeited and a penalty P is paid.

Assumption 1 (Types)**.** *The arrest probability depends on type: $p = p(\theta)$ with $p : [0, 1] \to (0, 1)$ continuously differentiable and $p'(\theta) < 0$. Higher-θ individuals face lower arrest risk — connections that provide protection, familiarity with criminal networks, lower detection. Types are distributed $\theta \sim F$ with continuous density $f_\theta > 0$; $\eta_i \sim G$ with continuous density $g > 0$ on an interval containing the thresholds arising below; θ and η_i are independent. Base wages satisfy $w_{I0} > w_{L0}$.*

C.2 Occupational and Educational Choices

C.2.1 Sector choice

Define the sector-optimized value of each schooling/field cell:

$$V_0(\theta) \equiv \max\big\{ w_{L0},\ (1 - p(\theta))\, w_{I0} - p(\theta)P \big\}, \tag{10}$$

$$V_N(\theta) \equiv \max\big\{ w_{L0} + \beta_L^N,\ (1 - p(\theta))\big(w_{I0} + \beta_I^N\big) - p(\theta)P \big\}, \tag{11}$$

$$V_C(\theta; \pi_L, \pi_I) \equiv \max\big\{ w_{L0} + \pi_L,\ (1 - p(\theta))\big(w_{I0} + \pi_I\big) - p(\theta)P \big\}. \tag{12}$$

Lemma 1 (Sector thresholds)**.** *(i) An uneducated individual works illegally if and only if*

$$p(\theta) < \bar{p}_0 \equiv \frac{w_{I0} - w_{L0}}{w_{I0} + P} \iff \theta > \bar{\theta}_0, \qquad p(\bar{\theta}_0) = \bar{p}_0. \tag{13}$$

(ii) An educated chemist works illegally if and only if

$$p(\theta) < \bar{p}_C \equiv \frac{(w_{I0} + \pi_I) - (w_{L0} + \pi_L)}{w_{I0} + \pi_I + P} \iff \theta > \bar{\theta}_C(\pi_L, \pi_I), \qquad p(\bar{\theta}_C) = \bar{p}_C, \tag{14}$$

which admits an interior solution only if $w_{I0} + \pi_I > w_{L0} + \pi_L$.

(iii) The threshold responds to prices as

$$\mu_L \equiv \frac{\partial \bar{\theta}_C}{\partial \pi_L} = \frac{1}{|p'(\bar{\theta}_C)|\,(w_{I0} + \pi_I + P)} > 0, \qquad \mu_I \equiv -\frac{\partial \bar{\theta}_C}{\partial \pi_I} = \frac{w_{L0} + \pi_L + P}{|p'(\bar{\theta}_C)|\,(w_{I0} + \pi_I + P)^2} > 0, \tag{15}$$

with $\mu_I/\mu_L = (w_{L0} + \pi_L + P)/(w_{I0} + \pi_I + P) < 1$ *whenever* $\bar{\theta}_C$ *is interior.*

Proof. (i) Illegal work dominates iff $(1-p)w_{I0} - pP > w_{L0}$, i.e. $w_{I0} - w_{L0} > p(w_{I0} + P)$; rearrange and use $p' < 0$. (ii) Illegal dominates iff $(1-p)(w_{I0} + \pi_I) - pP > w_{L0} + \pi_L$, i.e. $(w_{I0} + \pi_I) - (w_{L0} + \pi_L) > p\,(w_{I0} + \pi_I + P)$; if $w_{I0} + \pi_I \le w_{L0} + \pi_L$ the left side is nonpositive and no type chooses illegal chemistry work. (iii) Implicitly differentiate $p(\bar{\theta}_C)(w_{I0} + \pi_I + P) = (w_{I0} + \pi_I) - (w_{L0} + \pi_L)$ with respect to π_L and π_I:

$$\begin{aligned} p'(\bar{\theta}_C)\,\frac{\partial \bar{\theta}_C}{\partial \pi_L}\,(w_{I0} + \pi_I + P) = -1 \qquad &\Rightarrow \qquad \frac{\partial \bar{\theta}_C}{\partial \pi_L} = \frac{1}{|p'|(w_{I0} + \pi_I + P)}, \\ p'(\bar{\theta}_C)\,\frac{\partial \bar{\theta}_C}{\partial \pi_I}\,(w_{I0} + \pi_I + P) + p(\bar{\theta}_C) = 1 \qquad &\Rightarrow \qquad \frac{\partial \bar{\theta}_C}{\partial \pi_I} = -\frac{1 - p(\bar{\theta}_C)}{|p'|(w_{I0} + \pi_I + P)}. \end{aligned}$$

Substituting $1 - p(\bar{\theta}_C) = (w_{L0} + \pi_L + P)/(w_{I0} + \pi_I + P)$ from (14) gives (15); the ratio follows directly. □

An analogous threshold $\bar{\theta}_N$ governs educated non-chemists; its position relative to $\bar{\theta}_0$ depends on $\beta_I^N - \beta_L^N$ and plays no role in what follows.

Corollary 1 (Compensating differential)**.** *In any equilibrium in which some chemists work illegally, the marginal illegal chemist — type* $\bar{\theta}_C$ *— is indifferent between sectors:*

$$\big(1 - p(\bar{\theta}_C)\big)\big(w_{I0} + \pi_I\big) - p(\bar{\theta}_C)\,P \;=\; w_{L0} + \pi_L. \tag{16}$$

Given data on legal chemistry pay $(w_{L0} + \pi_L)$*, an arrest-risk estimate* $p(\bar{\theta}_C)$*, and a penalty value*

P, equation (16) backs out the total compensation of the marginal illegal chemist, $w_{I0} + \pi_I$, without any direct observation of illegal-sector pay.

Equation (16) is the first of three calibration devices the two-price structure delivers; the others appear in Section C.5.

C.2.2 Field and schooling choices

Conditional on educating, a student chooses C over N if and only if $V_C(\theta; \pi_L, \pi_I) - \eta_i > V_N(\theta)$, and educates at all if and only if $\max\{V_C - \eta_i,\, V_N\} - c > V_0(\theta)$.[24] Combining, an individual (θ, η_i) *educates and chooses chemistry* if and only if

$$\eta_i \;<\; \bar{\eta}(\theta; \pi_L, \pi_I) \;\equiv\; \min\Big\{ V_C - V_N,\;\; V_C - V_0 - c \Big\}, \tag{17}$$

where the first argument is the *field margin* and the second the *extensive (dropout) margin*.[25]

The object $V_C - V_0$, which governs the dropout margin for chemistry-bound students, has a three-region closed form:

$$V_C(\theta; \pi_L, \pi_I) - V_0(\theta) = \begin{cases} \pi_L, & \theta \le \bar{\theta}_0 \quad \text{(legal in both states)}, \\ \pi_L - R(\theta), & \bar{\theta}_0 < \theta \le \bar{\theta}_C \quad \text{(illegal if uneducated; legal if chemist)}, \\ (1 - p(\theta))\, \pi_I, & \theta > \bar{\theta}_C \quad \text{(illegal in both states)}, \end{cases} \tag{18}$$

where

$$R(\theta) \equiv (1 - p(\theta))\, w_{I0} - p(\theta) P - w_{L0} \tag{19}$$

is the *crime rent* of an uneducated type-θ individual. By Lemma 1, $R(\theta) > 0$ exactly on $\theta > \bar{\theta}_0$ and $R'(\theta) = |p'(\theta)|(w_{I0} + P) > 0$; the expression (18) is continuous at both thresholds.[26]

Equation (18) organizes the population into the three groups that the empirical analysis speaks to, and yields explicit decision rules for each.

[24] For a student who anticipates illegal work in either field, the field condition reads $(1 - p(\theta))(\pi_I - \beta_I^N) > \eta_i$: the factor $(1 - p(\theta))$ appears because arrest forfeits the wage while η_i is sunk at study time. This refines the condition $\beta_s^C - \beta_s^N > \eta_i$ used informally earlier; none of the results depend on the convention.

[25] If $V_N - c \le V_0$ and $\eta_i < V_C - V_0 - c$, then $V_C - \eta_i > V_0 + c \ge V_N$, so the field condition is implied and the two cases are exhaustive. If $V_N - c > V_0$, everyone of type θ educates and prices move only the field composition.

[26] At $\bar{\theta}_0$, $R(\bar{\theta}_0) = 0$. At $\bar{\theta}_C$, indifference (16) gives $(1 - p(\bar{\theta}_C))\, \pi_I = \pi_L - R(\bar{\theta}_C)$.

(a) High-θ (illegal-bound) youth. For $\theta > \bar{\theta}_C$ with $V_N - c \leq V_0$, education (in chemistry) is chosen if and only if

$$(1 - p(\theta))\,\pi_I > c + \eta_i \iff \theta > \theta_e(\eta_i, \pi_I), \qquad p\big(\theta_e\big) = 1 - \frac{c + \eta_i}{\pi_I}. \tag{20}$$

Implicit differentiation gives

$$\frac{\partial \theta_e}{\partial \pi_I} = \frac{(c + \eta_i)/\pi_I^2}{p'(\theta_e)} < 0 : \tag{21}$$

a higher illegal premium lowers the type threshold for schooling among crime-prone youth.Its empirical counterpart is that dropout falls in exposed areas within high-crime-propensity populations.

(b) Sector-switching (marginal) youth: the option value. Individuals with $\theta \in (\bar{\theta}_0, \bar{\theta}_C]$ work illegally if uneducated but legally if educated in chemistry: for them, schooling is a route *out of* illegal work.

Proposition 1 (Marginal types solved)**.** *For $\theta \in (\bar{\theta}_0, \bar{\theta}_C]$ with $V_N - c \leq V_0$, education is chosen if and only if*

$$\eta_i < \pi_L - c - R(\theta) : \tag{22}$$

the legal chemistry premium must cover the schooling cost plus the crime rent foregone by leaving illegal work. Equivalently, given η_i, the individual educates iff $\theta < \hat{\theta}(\eta_i, \pi_L)$ with $R(\hat{\theta}) = \pi_L - c - \eta_i$. Moreover: (i) the education rule of these types responds to π_L with slope one, the maximal response in the population; (ii) the group itself is an equilibrium object, since $\bar{\theta}_C$ moves with both prices by (15); (iii) their response has the same sign as that of groups (a) and (c).

Proof. On the region, (18) gives $V_C - V_0 = \pi_L - R(\theta)$, so the extensive-margin cutoff in (17) is $\bar{\eta} = \pi_L - R(\theta) - c$, which is (22); $R' > 0$ makes the equivalent θ-threshold well defined. (i) is immediate; (ii) is Lemma 1(iii); (iii) follows because every branch of (18) is increasing in (π_L, π_I). □

(c) Low-θ youth. For $\theta \leq \bar{\theta}_0$ with $\beta_L^N < c$ (the empirically relevant case, in which a dropout margin exists), education is chosen iff $\pi_L - \eta_i > c$, i.e. $\eta_i < \pi_L - c$ — a condition in which no illegal-sector object appears. (If $\beta_L^N \geq c$, all low-θ youth educate and π_L moves only their field composition.)

C.3 The Market for Chemistry Skill

C.3.1 Supply

Aggregating (17) over the population and splitting by sector of employment (Lemma 1) gives the supplies of chemistry-skilled labor to each sector:

$$H_L^{sup}(\pi_L, \pi_I) = \int_0^{\bar{\theta}_C} G\big(\bar{\eta}(\theta;\cdot)\big)\, dF(\theta), \qquad H_I^{sup}(\pi_L, \pi_I) = \int_{\bar{\theta}_C}^{1} G\big(\bar{\eta}(\theta;\cdot)\big)\, dF(\theta). \tag{23}$$

Lemma 2 (Supply responses)**.** *Define the entry margins and the sector-margin mass*

$$E_L \equiv \int_0^{\bar{\theta}_C} g(\bar{\eta})\, dF(\theta) \geq 0, \qquad E_I \equiv \int_{\bar{\theta}_C}^{1} g(\bar{\eta})\,(1-p(\theta))\, dF(\theta) \geq 0, \qquad m \equiv f_\theta(\bar{\theta}_C)\, G\big(\bar{\eta}(\bar{\theta}_C;\cdot)\big) \geq 0. \tag{24}$$

Then, applying Leibniz's rule to (23) *and using* (15)*,*

$$\begin{aligned} \frac{\partial H_L^{sup}}{\partial \pi_L} &= E_L + m\,\mu_L > 0, & \frac{\partial H_L^{sup}}{\partial \pi_I} &= -\,m\,\mu_I \leq 0, \\ \frac{\partial H_I^{sup}}{\partial \pi_L} &= -\,m\,\mu_L \leq 0, & \frac{\partial H_I^{sup}}{\partial \pi_I} &= E_I + m\,\mu_I > 0. \end{aligned} \tag{25}$$

Own-price responses combine an entry margin (E: new chemists drawn across the field or dropout cutoffs $\bar{\eta}$, with slope 1 for legal-bound types and $1-p(\theta)$ for illegal-bound types) and a sector margin ($m\mu$: existing chemists crossing $\bar{\theta}_C$); cross-price responses operate through the sector margin alone. The associated supply elasticities are

$$\varepsilon_{LL} \equiv \frac{\pi_L(E_L + m\mu_L)}{H_L^{sup}}, \quad \varepsilon_{LI} \equiv -\frac{\pi_I\, m\mu_I}{H_L^{sup}}, \quad \varepsilon_{IL} \equiv -\frac{\pi_L\, m\mu_L}{H_I^{sup}}, \quad \varepsilon_{II} \equiv \frac{\pi_I(E_I + m\mu_I)}{H_I^{sup}}. \tag{26}$$

Proof. For $\theta < \bar{\theta}_C$ the binding value V_C is the legal branch, so $\partial\bar{\eta}/\partial\pi_L = 1$ and $\partial\bar{\eta}/\partial\pi_I = 0$ (both arguments of the min in (17) shift with V_C, and V_N, V_0 are price-independent); for $\theta > \bar{\theta}_C$ the illegal branch binds, so $\partial\bar{\eta}/\partial\pi_I = 1 - p(\theta)$ and $\partial\bar{\eta}/\partial\pi_L = 0$. Differentiating (23) then yields interior terms E_L, E_I and boundary terms $\pm f_\theta(\bar{\theta}_C)G(\bar{\eta}(\bar{\theta}_C))\,\partial\bar{\theta}_C/\partial\pi$, which (15) evaluates to the $m\mu$ expressions. □

The entry margins map one-for-one into the paper's outcome variables: the components of E_L and E_I where the field margin binds are the compositional enrollment shift toward chemistry programs (Prediction 2), and the components where the dropout margin binds are the extensive-margin schooling response (Predictions 1 and 4). The sector margin m has no education counterpart

— it reallocates already-trained chemists — but it is what transmits illegal demand into the legal wage.

C.3.2 Demand

Assumption 2 (Iso-elastic sector demands)**.** *Demands for chemistry-skilled labor are*

$$H_L^d(\pi_L) = B\, D_L\, \pi_L^{-\sigma_L}, \qquad H_I^d(\pi_I) = A\, D_I\, \pi_I^{-\sigma_I}, \qquad D_L, D_I, \sigma_L, \sigma_I > 0. \tag{27}$$

China's 2019 fentanyl ban, by shifting illegal drug production synthetsis to Mexico, is an increase in the illegal shifter A in exposed commuting zones. The legal shifter B may respond to the same shock with pass-through $\phi \equiv d\ln B/d\ln A \geq 0$: $\phi > 0$ captures the demand-side linkages discussed in the mechanisms analysis — firm creation in precursor and equipment supply, agglomeration in chemical commerce, and quasi-legal activity supporting production.[27] *The baseline case is $\phi = 0$ (sectors only linked by pure labor-market competition); Section C.4 shows how the data can discipline ϕ.*

For compactness, write the demand slopes as $\delta_L \equiv \sigma_L H_L^d/\pi_L$ and $\delta_I \equiv \sigma_I H_I^d/\pi_I$.

C.3.3 Equilibrium

Both sectors recruit from the same local pool of chemistry-skilled workers, so the two premiums are determined jointly.

Proposition 2 (Existence and uniqueness)**.** *Suppose that, at every price pair,*

$$m\mu_I \;<\; E_L + m\mu_L + \delta_L \qquad \text{and} \qquad m\mu_L \;<\; E_I + m\mu_I + \delta_I. \tag{28}$$

Then under Assumptions 1–2 there exists a unique equilibrium price pair (π_L^*, π_I^*) solving

$$H_L^{sup}(\pi_L, \pi_I) = H_L^d(\pi_L), \qquad H_I^{sup}(\pi_L, \pi_I) = H_I^d(\pi_I). \tag{29}$$

The system exhibits gross substitutes: excess demand $Z_s \equiv H_s^d - H_s^{sup}$ is strictly decreasing in its own premium and strictly increasing in the other's.

Condition (28) requires each sector's clearing condition to respond more strongly to its own premium than to the other's. The first inequality holds automatically, since $\mu_I < \mu_L$ by Lemma 1(iii).

[27] Both demands can be microfounded by price-taking producers with decreasing returns to chemistry-skilled labor and iso-elastic output demand; σ_s then combines the output-demand elasticity and the labor share, and A collects the profitability of synthetic production per unit of chemistry skill. Only the constant-elasticity form is used below.

The second can fail only if the sector margin m is large relative to the illegal sector's own entry and demand responses — that is, only if a thick mass of chemists sits exactly at $\bar{\theta}_C$, so that the two sectors' workforces are near-perfect substitutes at the margin.

Proof. Gross substitutes. By (25) and $\delta_s > 0$, $\partial Z_L/\partial \pi_L = -(E_L + m\mu_L + \delta_L) < 0$ and $\partial Z_L/\partial \pi_I = m\mu_I > 0$; symmetrically for Z_I.

Existence. Fix π_I and vary π_L. As $\pi_L \to 0$, demand diverges under (27) while supply remains bounded. As $\pi_L \to \infty$, demand vanishes, while V_C grows without bound and V_N, V_0 do not, so both arguments of the minimum in (17) diverge, $G(\bar{\eta}) \to 1$ at every type, and supply increases to a strictly positive limit. With own-price monotonicity, market L therefore clears at a unique price for each π_I, continuous and increasing in π_I. The same holds for market I: there the range of integration in (23) widens as π_I grows, by (15), reinforcing rather than offsetting the effect through $\bar{\eta}$. Composing the two clearing functions maps a compact interval into itself, and the composition is continuous, so a fixed point exists by the intermediate value theorem.

Uniqueness. Let (π_L, π_I) and (π'_L, π'_I) both solve (29). Subtracting the clearing conditions and applying the mean value theorem to each separately, there are intermediate price pairs at which

$$(E_L + m\mu_L + \delta_L)(\pi'_L - \pi_L) = m\mu_I\,(\pi'_I - \pi_I), \qquad (E_I + m\mu_I + \delta_I)(\pi'_I - \pi_I) = m\mu_L\,(\pi'_L - \pi_L).$$

By gross substitutes the two differences share a sign; suppose both are positive, the other case being symmetric. The first equation with (28) gives $\pi'_L - \pi_L < \pi'_I - \pi_I$, and the second gives the reverse inequality. Both differences therefore vanish. □

The equilibrium premiums inherit the two-margin structure of Lemma 2. When $m = 0$ the clearing conditions decouple and each premium is determined by its own sector alone; the model then reduces to two independent one-price markets, and the transmission mechanism studied below is shut. Everything that follows turns on $m > 0$.

C.4 Effects of Skill-Biased Change in Illegal Markets

Consider the shock $d \ln A > 0$. Totally differentiating (29) and collecting terms,

$$\underbrace{\begin{pmatrix} E_L + m\mu_L + \delta_L & -m\mu_I \\ -m\mu_L & E_I + m\mu_I + \delta_I \end{pmatrix}}_{J} \begin{pmatrix} d\pi_L \\ d\pi_I \end{pmatrix} = \begin{pmatrix} \phi\, H_L^d \\ H_I^d \end{pmatrix} d \ln A. \tag{30}$$

Lemma 3 (Regularity)**.** $\det J > 0$ *always. Expanding,*

$$\det J = (E_L + \delta_L)(E_I + \delta_I) + (E_L + \delta_L)\, m\mu_I + (E_I + \delta_I)\, m\mu_L \;>\; 0, \tag{31}$$

because the cross terms $m^2\mu_L\mu_I$ *cancel exactly: the sector margin only reallocates chemists between sectors and does not affect* $\det J$*.*

Proof. $\det J = (E_L + m\mu_L + \delta_L)(E_I + m\mu_I + \delta_I) - m^2\mu_L\mu_I$. Multiplying out, the term $m\mu_L \cdot m\mu_I$ cancels against $-m^2\mu_L\mu_I$, leaving (31), a sum of nonnegative terms that is strictly positive since $\delta_L, \delta_I > 0$. □

Proposition 3 (Both premiums rise; the legal premium rises by less)**.** *By Cramer's rule applied to* (30)*,*

$$\frac{d\pi_I}{d\ln A} = \frac{(E_L + m\mu_L + \delta_L)\, H_I^d + m\mu_L\, \phi H_L^d}{\det J} \;>\; 0, \tag{32}$$

$$\frac{d\pi_L}{d\ln A} = \frac{m\mu_I\, H_I^d + (E_I + m\mu_I + \delta_I)\, \phi H_L^d}{\det J} \;>\; 0 \quad \textit{whenever } m > 0 \textit{ or } \phi > 0. \tag{33}$$

In the pure-competition case $\phi = 0$*,*

$$0 \;<\; \frac{d\pi_L}{d\pi_I} = \frac{m\mu_I}{E_L + m\mu_L + \delta_L} \;<\; 1 : \tag{34}$$

the legal premium rises, but by less than the illegal premium, since $m\mu_I < m\mu_L \le E_L + m\mu_L + \delta_L$ *by Lemma 1(iii).*

Proof. Cramer's rule with Lemma 3; every entry in the numerators is nonnegative, and (33) is strictly positive when either transmission channel ($m > 0$: sector competition; $\phi > 0$: demand linkages) is open. The ratio (34) divides (33) by (32) at $\phi = 0$ and uses $\mu_I < \mu_L$. □

Prediction 1 (Legal Sector Spillovers — derived)**.** *An outward shift in illegal demand for chemistry skill raises the chemistry premium in the* legal *sector (equation* (33)*). The transmission requires no appeal to assumed spillover channels: because both sectors recruit from the same local pool of chemists, illegal expansion bids up the illegal premium, pulls marginal chemists across* $\bar{\theta}_C$*, and the induced scarcity raises the legal premium. Demand-side linkages (*$\phi > 0$*) amplify but are not necessary for the wage result.*

Corollary 2 (Employment separates poaching from linkages)**.** *Legal-sector chemistry employment*

responds as $d \ln H_L^d = \phi \, d \ln A - \sigma_L \, d \ln \pi_L$*, which is positive if and only if*

$$\phi \;>\; \underline{\phi} \equiv \frac{\sigma_L \, |\varepsilon_{LI}|}{\varepsilon_{LL} \, (\varepsilon_{II} + \sigma_I) - \varepsilon_{LI} \, \varepsilon_{IL}} \;\geq\; 0. \tag{35}$$

Under pure poaching ($\phi = 0 < \underline{\phi}$ *when* $m > 0$)*, legal chemistry* wages *rise while legal chemistry* employment *falls, as the illegal sector draws chemists away along the legal demand curve. The joint empirical finding that exposure raises both the legal chemistry premium and legal chemistry-related employment therefore implies* $\phi > \underline{\phi}$*: the boom shifted legal demand outward, consistent with the input-linkage and firm-creation channels examined in Section 8.*

Proof. Solve (30) in elasticity form. Dividing the two rows by H_L^{sup} and H_I^{sup} and using (26), the system is $(\varepsilon_{LL}+\sigma_L)\, x + \varepsilon_{LI}\, y = \phi a$ and $\varepsilon_{IL}\, x + (\varepsilon_{II}+\sigma_I)\, y = a$, where $x \equiv d \ln \pi_L$, $y \equiv d \ln \pi_I$, $a \equiv d \ln A$. Cramer's rule gives $x = a\,[\phi(\varepsilon_{II}+\sigma_I) + |\varepsilon_{LI}|]/\widetilde{D}$ with $\widetilde{D} \equiv (\varepsilon_{LL}+\sigma_L)(\varepsilon_{II}+\sigma_I) - \varepsilon_{LI}\varepsilon_{IL} > 0$ (the elasticity form of Lemma 3). Substituting into $d \ln H_L^d = \phi a - \sigma_L x$ and collecting terms yields positivity iff $\phi\,[\varepsilon_{LL}(\varepsilon_{II} + \sigma_I) - \varepsilon_{LI}\varepsilon_{IL}] > \sigma_L |\varepsilon_{LI}|$, which is (35). □

The education predictions now follow from Proposition 3 and the decision rules of Section C.2.

Prediction 2 (Direct Effect on Educational Attainment — derived)**.** *The shock raises* π_I *(32), which lowers the education threshold* $\theta_e(\eta_i, \pi_I)$ *for illegal-bound youth (21): educational investment rises among individuals with* $\theta > \theta_e$*, i.e. with sufficiently high propensity for illegal work, where "sufficiently high" is the explicit cutoff in (20).*

This prediction reverses the standard crime-economics result. In the traditional framework, criminal opportunities reduce educational investment by raising w_{I0} — the return to uneducated illegal labor — which increases the opportunity cost of schooling. Skill-biased change in illegal markets works through a different channel: it bids up the price of chemistry-intensive education, making schooling more valuable for those who anticipate illegal-sector work.

However, the expansion of synthetic drug production may also increase w_{I0} for some individuals. Laboratory operations require not only chemists but also support personnel — lookouts, security, transportation, and other tasks that criminal organizations typically assign to young, uneducated recruits. If laboratory expansion increases demand for these roles, the opportunity cost of schooling rises for youth already inclined toward criminal involvement, potentially offsetting the skill-biased effect. In the notation above, a rise in w_{I0} raises $R(\theta)$ and $V_0(\theta)$ for $\theta > \bar{\theta}_0$, tightening the education conditions (20) and (22) even as the higher premiums relax them. The net effect on educational attainment is thus theoretically ambiguous for this group, depending on whether the increased returns to chemistry skills outweigh the increased returns to unskilled criminal labor.

Because unskilled illegal labor is a distinct market from chemistry skill, this ambiguity does not contaminate the market-clearing results for (π_L^*, π_I^*).

Prediction 3 (Composition of Education — derived)**.** *Because the shock is skill-specific, both premiums rise while β_L^N and β_I^N are unchanged, so the field cutoff $V_C - V_N$ rises at every θ (Lemma 2). The share of educated individuals choosing chemistry, $G(V_C - V_N)$, increases; enrollment in chemistry-intensive programs rises; non-chemistry enrollment weakly falls as marginal students reallocate; and total enrollment changes only through the extensive margin (Predictions 2 and 4). The effect is primarily* compositional.

Prediction 4 (Indirect Effect on Educational Attainment — derived)**.** *Educational investment rises among individuals with low θ (no propensity for illegal work). For $\theta \leq \bar{\theta}_0$ with $\beta_L^N < c$, the education share is $G(\pi_L^* - c)$, and*

$$\frac{d\, G(\pi_L^* - c)}{d \ln A} = g(\pi_L^* - c)\, \frac{d\pi_L}{d \ln A} \;>\; 0 \tag{36}$$

by (33). Nothing in the low-θ decision rule references the illegal sector: the effect operates only through the equilibrium legal premium.

Predictions 2 and 4 both imply increased educational attainment, operating through the illegal and legal premium respectively. The model does not take a stand on the relative magnitudes of the three groups' responses — that depends on the local distribution F — and the empirical analysis accordingly interprets the education effects as their sum.

C.5 Magnitudes

A natural objection to the mechanism is that clandestine synthetic-drug production employs few chemists relative to the local labor market, so its demand could not plausibly move premiums or education outcomes. The market-clearing block shows this only depends on elasticities and can that even a small illegal sector can move the chemistry premium when the supply of chemists is inelastic.

Proposition 4 (Inelastic supply amplifies)**.** *In elasticity form (proof of Corollary 2), the premium responses at $\phi = 0$ are*

$$\frac{d \ln \pi_L}{d \ln A} = \frac{|\varepsilon_{LI}|}{\widetilde{D}}, \qquad \frac{d \ln \pi_I}{d \ln A} = \frac{\varepsilon_{LL} + \sigma_L}{\widetilde{D}}, \qquad \widetilde{D} = (\varepsilon_{LL} + \sigma_L)(\varepsilon_{II} + \sigma_I) - \varepsilon_{LI}\, \varepsilon_{IL} > 0, \tag{37}$$

and both are decreasing in the entry margins E_L, E_I (which enter $\varepsilon_{LL}, \varepsilon_{II}$ but not the cross elasticities). With slow entry, which is empirically relevant since producing an additional chemistry-trained worker requires years of schooling, amplifies the price response to any given demand shift.

Proof. Equation (37) restates the Cramer solution. The entry margins enter only the own elasticities: E_L raises ε_{LL} and E_I raises ε_{II}, while ε_{LI} and ε_{IL} depend on the sector margin alone (26). Differentiating: $\partial(d\ln\pi_L/d\ln A)/\partial\varepsilon_{LL} = -|\varepsilon_{LI}|(\varepsilon_{II}+\sigma_I)/\widetilde{D}^2 < 0$; $\partial(d\ln\pi_I/d\ln A)/\partial\varepsilon_{LL} = [\widetilde{D} - (\varepsilon_{LL}+\sigma_L)(\varepsilon_{II}+\sigma_I)]/\widetilde{D}^2 = -\varepsilon_{LI}\varepsilon_{IL}/\widetilde{D}^2 < 0$ (the product $\varepsilon_{LI}\varepsilon_{IL}$ of two negative cross elasticities is positive); $\partial(d\ln\pi_L/d\ln A)/\partial\varepsilon_{II} = -|\varepsilon_{LI}|(\varepsilon_{LL}+\sigma_L)/\widetilde{D}^2 < 0$; and $\partial(d\ln\pi_I/d\ln A)/\partial\varepsilon_{II} = -(\varepsilon_{LL}+\sigma_L)^2/\widetilde{D}^2 < 0$. Both premium responses are therefore strictly decreasing in both entry margins. □

Corollary 3 (Fixed-stock of chemists in the short run)**.** *Over horizons too short for schooling to add chemists ($E_L = E_I = 0$), the total stock is fixed and every chemist the illegal sector hires is a chemist the legal sector loses: $dH_L = -\,dH_I$. Moving along the legal demand curve,*

$$d\ln\pi_L = \frac{1}{\sigma_L}\frac{dH_I}{H_L}, \qquad \textit{equivalently} \qquad \frac{dH_I}{H_L} = \sigma_L\,\widehat{d\ln\pi_L}, \tag{38}$$

where $\widehat{d\ln\pi_L}$ is the estimated exposure effect on the legal chemistry premium. A second short-run mapping follows from the first equation of (37):

$$\widehat{d\ln\pi_I} = \frac{\varepsilon_{LL}+\sigma_L}{|\varepsilon_{LI}|}\,\widehat{d\ln\pi_L}, \tag{39}$$

which identifies the (unobserved) illegal-premium response from the (estimated) legal one; combined with Corollary 1 it yields the implied illegal chemist compensation in levels.

Proof. With $E_L = E_I = 0$ and full employment of the fixed stock, market clearing forces $dH_L = -dH_I$; the legal demand curve (27) gives $d\ln H_L = -\sigma_L\,d\ln\pi_L$, and substituting delivers (38). Equation (39) rearranges the legal-market row of the log-linear system at $\phi = 0$: $(\varepsilon_{LL}+\sigma_L)\,d\ln\pi_L = |\varepsilon_{LI}|\,d\ln\pi_I$. □

Equation (38) can be used to identify the movements in the illega sector from the paper's difference-in-differences estimates. The legal premium response equals the illegal sector's incremental hiring *as a share of local legal chemistry employment*, scaled by the inverse legal demand elasticity. Equivalently, the DiD estimate of the legal chemistry premium, multiplied by σ_L, is the model-implied poaching share. The plausibility of the chemistry skill mechanism is thereby an elasticity question — how inelastic are legal demand and short-run supply — rather than the size of the

illegal sector: with specialized legal demand (σ_L low) and an educational pipeline that takes years to respond ($E \approx 0$ over the study horizon), an illegal sector hiring a few percent of local chemists moves the legal premium by several percent, with the education responses of Section C.4 following. The same inelasticity implies that quantity (enrollment) responses build gradually, which is consistent with the growing event-study profiles in Section 6.

Two caveats discipline the use of these results. First, the calibration devices (16), (38), and (39) are *plausibility and consistency* checks: they show that the estimated wage and education responses are consistent with the skill-demand mechanism under defensible elasticities but do not identify the mechanism, the causal identification comes from the difference-in-differences design in Section 6. Second, the fixed-stock case deliberately shuts down the education margins that are the paper's outcomes; it is a device for interpreting the wage estimates, and the full model (37) governs the medium run in which enrollment responds.

D Assigning Tertiary Enrollment to Origin Commuting Zones

University enrollment is recorded where each campus is located, with no information on where its students come from. Treatment, however, is defined in the students' home commuting zone. This appendix describes how I assign enrollment to origin commuting zones, and the assumptions under which this imputation measures changes in total enrollment, study destinations, and enrollment by field of study. Appendix H.6 reports the corresponding robustness checks.

D.1 Construction of the migration matrix

The weights come from the moves of university-age students between 2005 and 2010 recorded in the 2010 Population Census. I use the long-form questionnaire of the census, a 10% sample, with its expansion factors. The sample consists of individuals aged 19 to 24, the ages of university students, who report attending school. For each individual I observe the commuting zone of residence in 2010, which I treat as the destination d, and the commuting zone of residence in 2005, which I treat as the origin o. Let $M_{o,d}$ denote the weighted number of students who lived in o in 2005 and in d in 2010. The weight used in the imputation is

$$P_{d,o} = \frac{M_{o,d}}{\sum_{o'} M_{o',d}}, \tag{40}$$

the share of the students living in destination d whose origin is o. These shares sum to one over origins for each destination. Imputed enrollment from origin o in year t is

$$\tilde{Y}_{o,t} = \sum_{d} P_{d,o}\, Y_{d,t}, \tag{41}$$

where $Y_{d,t}$ is ANUIES enrollment in destination d. Students who lived abroad in 2005 or whose 2005 municipality is not reported are part of each destination's total but are not assigned to any origin. Together with enrollment assigned to commuting zones outside the estimation sample, this unassigned enrollment amounts to 1.9% of enrollment before 2019. Figure B.2 describes the distribution of the weights. The same matrix is used for all fields of study and gender; Appendix H.6 reports estimates with separate matrices by sex.

Most students study in their own commuting zone, and few cross treatment status. Table D.1 summarizes student flows between treated and control commuting zones. In 2010, 96.1% of students from treated origins and 92.0% of students from control origins studied in their own commuting zone. Only 1.4% of students from treated origins study in a control commuting zone, and 3.1%

of students from control origins study in a treated one. Among students who move, however, more than a third cross treatment status. On the destination side, 2.8% of the students living in treated commuting zones come from control origins.

Table D.1: Student flows between treated and control commuting zones, 2005–2010

	By origin		By destination	
	Treated	Control	Treated	Control
Enrolled in own commuting zone	96.1	92.0	93.3	92.2
Crossing treatment status, all students	1.4	3.1	2.8	1.5
Crossing treatment status, movers only	35.9	38.7	41.8	19.2

Notes: Percentages of students aged 19 to 24 attending school in 2010, weighted by census expansion factors. Origin is the commuting zone of residence in 2005 and destination the commuting zone of residence in 2010. The "by origin" columns divide by all students from origins of each treatment status; the "by destination" columns divide by all students living in destinations of each treatment status. "Crossing treatment status" counts students whose origin and destination differ in treatment status. "Movers only" divides by students who changed commuting zone. Students who lived abroad or whose 2005 municipality is unknown are excluded from the origin columns and counted as non-local, non-crossing students in the destination columns. Treated commuting zones are those with a methamphetamine laboratory seized before 2015.
Source: Mexican Population Census 2010, long-form questionnaire; author's calculations.

D.2 Limitations of the imputation approach

The imputation has to be judged against the choices young people make at the university level. The research question concerns how young people adjust their education to the shock, and at the university level they make three choices: whether to enroll, where to study, and what to study. Each choice can respond to the shock. A higher return to some skills may lead more young people to enroll, or more of them to choose the fields that teach those skills, such as chemistry. The shock can also change where they study if, for example, new programs in treated areas attract students from elsewhere, or if students seek distance from areas with criminal activity and decide to study elsewhere.

Let $E_{o,t}$ be the number of university students from origin o in year t and $s_{o,d,t}$ the share of them enrolled in destination d. Enrollment in university from o in d is

$$N^{f}_{o,d,t} = E_{o,t}\, s_{o,d,t} \tag{42}$$

The data records enrollment by field in each destination, $Y_{d,t} = \sum_o N_{o,d,t}$, and the imputation assigns it to origins with the 2010 weights $P_{d,o}$. This recovers the change in enrollment from each

origin under the following assumption:

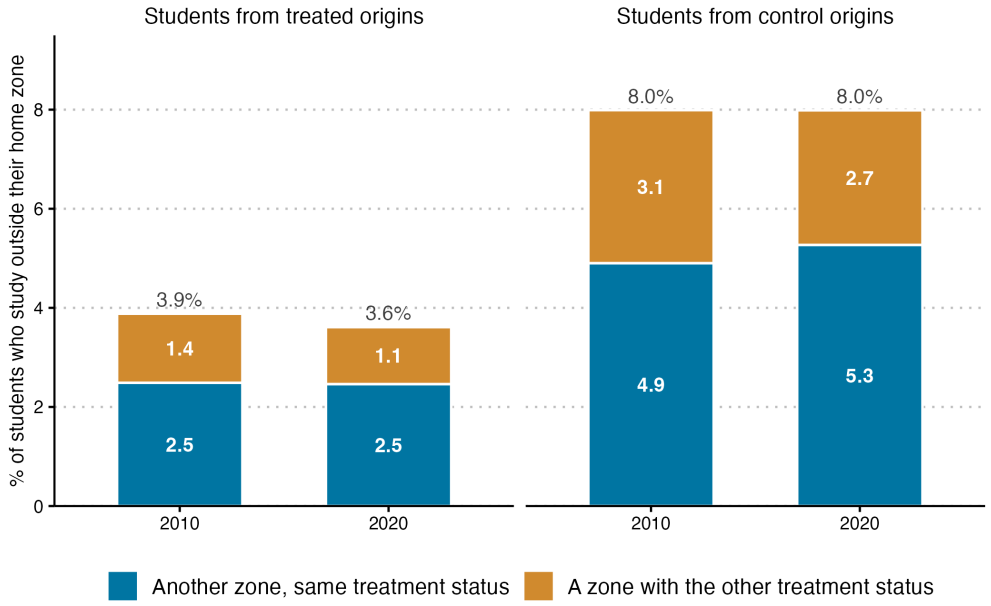
> *Assumption 1.* The share of an origin's students in each destination is the same as in 2010: $s_{o,d,t} = s_{o,d}$.

Fixed destination shares. Assumption 1 allows the number of students from each origin and their choice of field to change, but rules out changes in where they study.

Between the 2010 and 2020 censuses, the share of students who study at home and the split between treated and control destinations changed little. The 2020 Census, measured after the ban, allows a direct comparison with the 2010 matrix. With only two censuses, I cannot formally test whether exposure to synthetic drug production changed where students study, but the comparison bounds the contribution of such changes and shows whether they could affect the interpretation of the estimates. Figure D.1 summarizes the differences between 2010 and 2020. Most students study in their home commuting zone in both censuses: 96.1% of students from treated origins in 2010 and 96.4% in 2020, and 92.0% of students from control origins in both years. The share who study in a zone with the other treatment status fell between the two censuses, from 1.4% to 1.1% for students from treated origins and from 3.1% to 2.7% for students from control origins.

The matrix changes more at the level of individual destinations. Table D.2 measures the reallocation between censuses: the share of an origin's students who would have to change destination to go from the 2010 shares to the 2020 shares. Among the 405 origins with at least 500 students in both censuses, which account for 94% of students, this share is 2.5% for treated origins and 6.5% for control origins. Part of it reflects sampling, since both censuses are samples. Most of the reallocation takes place between zones with the same treatment status, which has little effect on the comparison between treated and control origins. Moves to or from zones with the other treatment status account for 31% of the reallocation for treated origins and 26% for control origins, and they largely offset each other. Assumption 1 therefore holds closely for the share of students who study at home and for the split between treated and control destinations.

Figure D.1: University students destinations outside of their home commuting zone, 2010 and 2020 censuses



Notes: Percentage of students aged 19 to 24 attending school who study outside the commuting zone where they lived five years earlier, by the treatment status of the commuting zone where they study, weighted by census expansion factors. The remaining students study in their home commuting zone. Treated commuting zones are those with a methamphetamine laboratory seized before 2015.
Source: Mexican Population Censuses 2010 and 2020, long-form questionnaires; author's calculations.

Table D.2: Changes in where students study between the 2010 and 2020 censuses

	Share of students (%)			Reallocation (pp)	
	2010	2020	Change	By destination type	Share across status
Students from treated origins					
Own commuting zone	96.1	96.4	0.27	0.57	
Another zone, same status	2.5	2.5	-0.03	1.16	
Zone with the other status	1.4	1.1	-0.25	0.77	
Total reallocation				2.49	31%
Students from control origins					
Own commuting zone	92.0	92.0	0.00	1.85	
Another zone, same status	4.9	5.3	0.37	2.96	
Zone with the other status	3.1	2.7	-0.37	1.71	
Total reallocation				6.51	26%

Notes: Students aged 19 to 24 attending school, weighted by census expansion factors. The first three columns give the share of students from treated or control origins who study in each type of destination, for all origins. The reallocation columns use origins with at least 500 students in both censuses, which account for 94% of students. For each origin, reallocation is half the sum, over destinations, of the absolute change in the share of the origin's students who study there; it equals the share of students who would have to change destination to go from the 2010 to the 2020 shares. "By destination type" gives the part of this reallocation that involves each type of destination, averaged over origins with 2010 students as weights, and "Total reallocation" their sum. "Share across status" is the part that involves zones with the other treatment status, as a percentage of the total.
Source: Mexican Population Censuses 2010 and 2020, long-form questionnaires; author's calculations.

Same origins in every field. Total enrollment from an origin is assigned with weights built from all students. To observe enrollment by field, the imputation uses the same weights $P_{d,o}$ for every field, which requires a second assumption:

> *Assumption 2.* Within each destination, students in every field come from the same mix of origins: in 2010, a student of field f in destination d came from origin o with the same probability $P_{d,o}$ as any other student in d.

Assumption 2 can fail for fields offered in only a few places. The assumption concerns the matrix, so it can fail even if nothing changes over time. Specialized programs, for example, are concentrated in large universities and may draw students from a wider set of origins than programs offered everywhere. In that case the imputation assigns enrollment in that field to the wrong mix of origins, and a change in enrollment in the field at a destination is attributed to origins in proportion to the destination's students in all fields.

The available data do not allow a systematic test of this assumption, but the campus-level estimates show that the results do not depend on it. The administrative data do not record where students come from. The census does, but a separate matrix for each field would rest on very few sampled students in most origin–destination pairs. The campus-level estimates in Appendix H.6 do not use the matrix, and they are nearly identical to the origin estimates.

Attenuation factor Under both assumptions, the imputation recovers changes in whether to enroll and in what to study, up to an attenuation factor.

The derivation applies to enrollment in any field. Suppose both assumptions hold and that, after the ban, enrollment in a field from every treated origin grows by a factor $1+g$, because more young people enroll, because more of them choose that field, or both. Enrollment from control origins is unchanged: $E_{o,t} = E_{o,0}(1+gD_o)$, where $E_{o,t}$ now denotes enrollment in the field and D_o indicates a treated origin. Let

$$T_d = \sum_{o:\, D_o=1} P_{d,o} \tag{43}$$

be the share of destination d's 2010 students who come from treated origins. Destination enrollment becomes $Y_{d,t} = Y_{d,0}(1+gT_d)$, and imputed enrollment becomes $\tilde{Y}_{o,t} = \sum_d N_{o,d,0}(1+gT_d)$. Summing over the origins in treatment group $D \in \{0,1\}$,

$$\frac{\sum_{o:\, D_o=D} \tilde{Y}_{o,t}}{\sum_{o:\, D_o=D} \tilde{Y}_{o,0}} = 1 + g\,\bar{T}_D, \qquad \bar{T}_D = \frac{\sum_{o:\, D_o=D} \sum_d N_{o,d,0}\, T_d}{\sum_{o:\, D_o=D} \sum_d N_{o,d,0}}. \tag{44}$$

$\bar{T}_D$ is the average, over students from group D, of the share of their classmates who come from treated origins. A Poisson regression with commuting zone and year fixed effects compares the growth of these group totals, so the estimated effect is

$$\hat{\beta} = \ln(1+g\bar{T}_1) - \ln(1+g\bar{T}_0) \approx \lambda \ln(1+g), \qquad \lambda = \bar{T}_1 - \bar{T}_0. \tag{45}$$

The imputation keeps the sign of the effect and shrinks it toward zero. The true effect is $\ln(1+g)$. Students from treated origins study mostly alongside other students from treated origins, so $\bar{T}_1$ is close to one, $\bar{T}_0$ is close to zero, and λ lies between zero and one. The estimate is smaller than the true effect because part of the increase at shared destinations is assigned to control origins. The same factor shrinks negative effects. Since λ depends only on the matrix and pre-period enrollment, I compute it directly in Appendix H.6.

The robustness checks in Appendix H.6 show that the results do not depend on the matrix. The estimates are nearly identical with separate matrices by sex, with a matrix built from the 2020

Census, and with enrollment measured where each campus is located, without any assignment to origins. It also reports the attenuation factor λ, and it compares the estimates with school attendance at tertiary age in ENOE, which does not use the matrix.

E Machine Learning Prediction of Latent Production

This appendix provides details on the machine learning approach used to predict latent synthetic drug production capacity. The methodology builds on the growing literature applying machine learning to policy problems (Athey and Imbens, 2019,0; Kleinberg et al., 2015; Mullainathan and Spiess, 2017). Similar approaches have been used to detect concealed illegal activity in other contexts like tax fraud detection (Battagliola and others, 2024; de Roux et al., 2018). Closely related, Campedelli et al. (2024) use machine learning to predict mafia infiltration of Italian municipalities—observed only when the central state dismisses an infiltrated city council—and use the resulting index to flag municipalities not yet detected by enforcement; as with laboratory seizures, council dismissals reflect enforcement effort and understate true presence. The common challenge in these applications is that observed violations reflect both underlying illegal activity and enforcement effort, motivating prediction-based approaches that identify likely violators based on characteristics associated with the activity itself rather than detection.

E.1 Motivation

Observed laboratory seizures D_i reflect the intersection of latent production D_i^* and enforcement:

$$D_i = D_i^* \cdot E_i \tag{46}$$

where E_i indicates The model is trained on observed seizures D_i, recovering $\hat{P}(D_i = 1 \mid X_i)$. This ranks commuting zones by their production potential independent of enforcement, $\hat{P}(D_i^* = 1|X_i)$ under the assumption that, conditional on X_i, detection is independent of latent production. This is, enforcement targets areas The goal is to predict $\hat{P}(D_i^* = 1|X_i)$—production potential independent of detection—to identify CZs that may have evaded enforcement despite having production capacity.

E.2 Features

All features are measured in 2015 or earlier to ensure they are pre-determined relative to the treatment period. I focus on characteristics that predict where production is viable rather than where authorities detect it.

Economic Fundamentals.

- *Market access:* Distance to U.S. border (primary export market), road network connectivity

- *Input access:* Distance to international ports on Pacific coast (precursor import routes), presence of universities with chemistry programs (ANUIES), number of chemical and fentanyl-related firms (Economic Census)
- *Demographics:* Population by age and gender, schooling levels, migration patterns (Population Census)

Concealment and Competition.

- *Geography and climate:* 19 bioclimatic variables from WorldClim, elevation, slope, terrain ruggedness
- *Illegal activities:* Cartel presence (Sobrino, 2019), other drug seizures (SEDENA), agricultural suitability for drug crops (Daniele et al., 2023)
- *Governance:* Distance to state capital, distance to Mexico City

I deliberately exclude variables related to enforcement intensity, police presence, or government capacity, which would predict detection rather than production.

E.3 Models

I estimate three models:

- *Random Forest:* Grows many deep trees in parallel, each on a bootstrap resample and choosing splits from a random subset of features.
- *XGBoost:* Grows trees sequentially, with each tree correcting errors from previous iterations. Often achieves higher accuracy but has a greater risk of overfitting.
- *Ensemble:* Averages predictions from Random Forest and XGBoost.

E.4 Spatial Cross-Validation

Standard cross-validation risks learning where authorities detect labs rather than where labs exist, as nearby CZs may share enforcement regimes. I implement spatial cross-validation to force generalization across enforcement environments.

I partition Mexico into five geographic folds using spatial clustering (Figure E.1). In each iteration, the model trains on four folds and predicts on the held-out fold; Figure E.2 illustrates one such

training–test split. This ensures that predictions for each CZ are based on patterns learned from geographically distant regions with potentially different enforcement intensity. Commuting zones are grouped into 25 spatial clusters, which are assigned whole to folds so that each cluster is never predicted using observations from itself. Under random cross-validation the median held-out zone has a training zone 15 km away, and 76% have one within 25 km — performance would substantially reflect spatial interpolation. Under the spatial folds the median distance rises to 62 km (78 km among seizure zones) and only 14% retain a training neighbour within 25 km. Still, clusters are distributed across folds so that each fold contains some zones with seizures, which limits how far apart training and test areas can be. The aim is to prevent predictions from leaning on immediate neighbours.

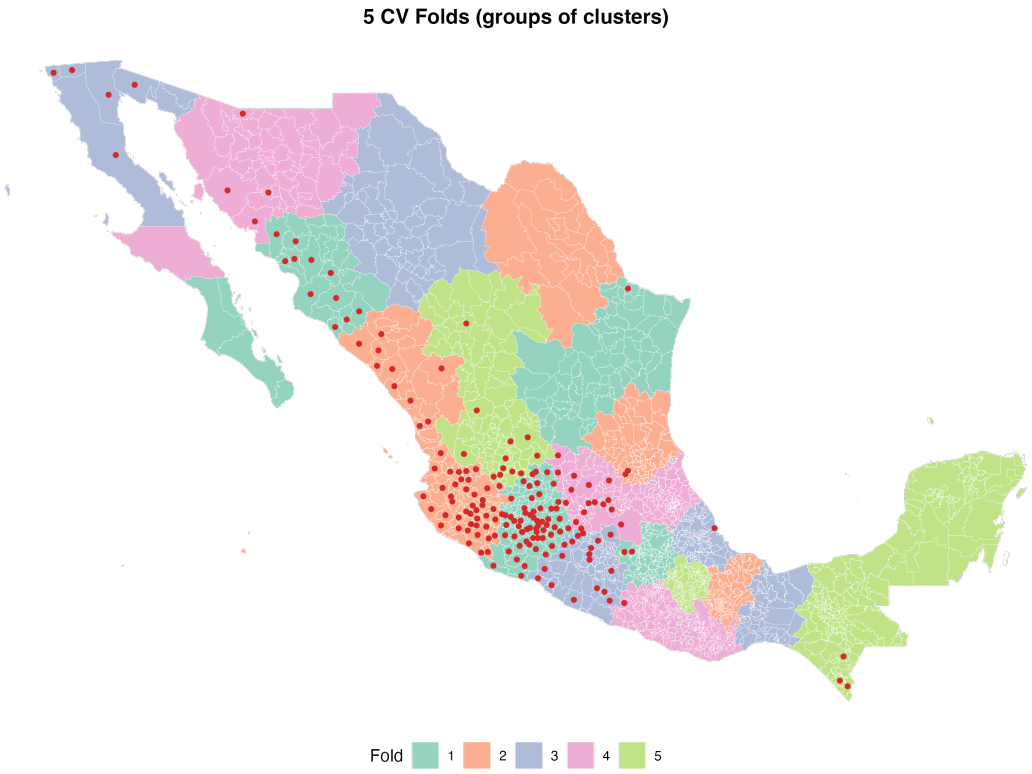
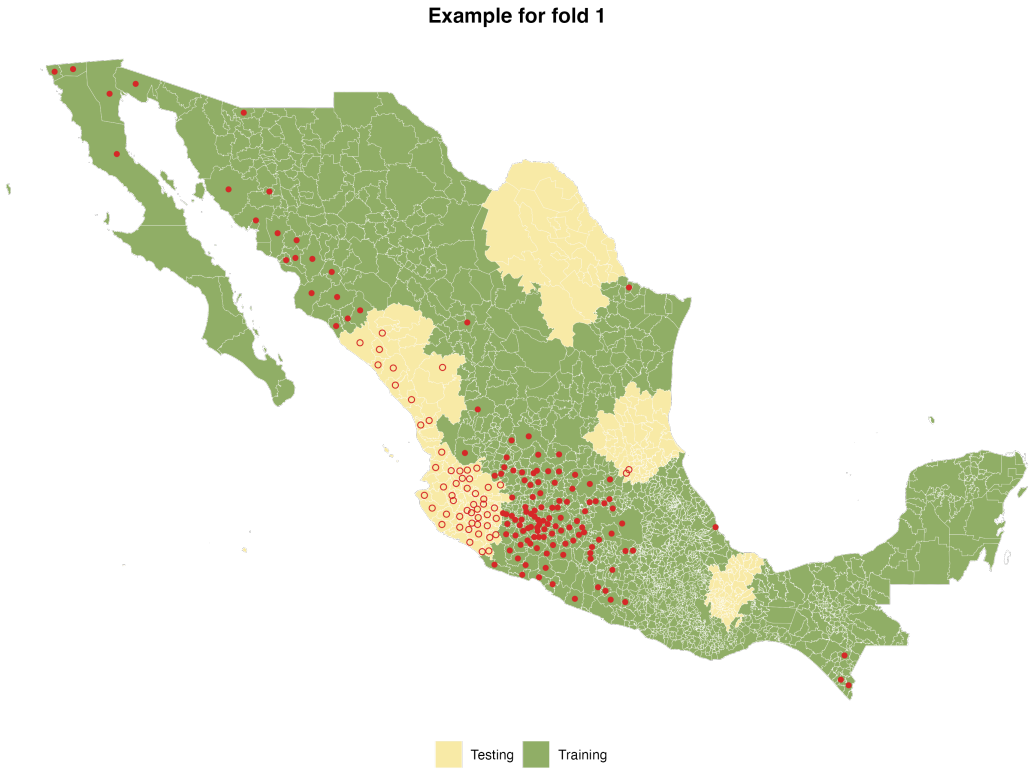


Figure E.1: Spatial Cross-validation folds

Notes: The map shows the five spatially-clustered cross-validation folds used to train and evaluate the machine-learning models that predict latent laboratory presence. In each iteration the model trains on four folds and forms predictions for the held-out fold, so that predictions for a commuting zone rely on geographically distant regions; this guards against spatial leakage between training and test sets.

↩ Return to text

Figure E.2: Example of training and testing samples within a fold



Notes: Illustrative training and test split for one spatial cross-validation fold. Commuting zones in the held-out fold (test) are predicted using a model trained on the remaining folds (training).

⟷ Return to text

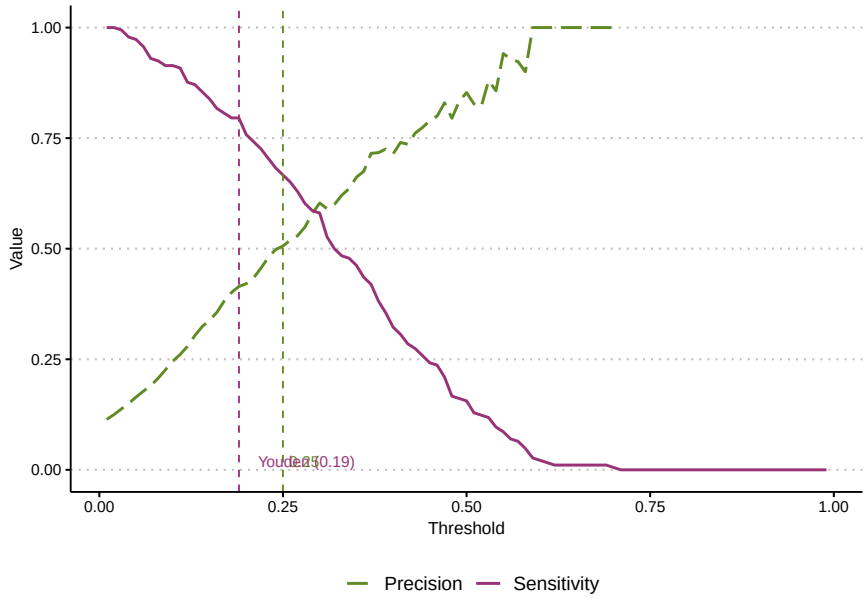
E.5 Classification Threshold

Translating the continuous predicted probability $\hat{P}(D_i^* = 1 \mid X_i)$ into a binary treatment requires a classification threshold τ. There is no single "correct" threshold: because true production is unobserved and observed seizures conflate production with enforcement, the usual recall and precision criteria are themselves measured against an imperfect label. Figure E.3 plots this trade-off, showing recall (sensitivity) and precision against τ when the baseline seizure indicator is taken as the (imperfect) ground truth. As τ falls, the model flags more commuting zones, raising recall at the cost of precision.

A lower threshold has the appeal of capturing potential producers that enforcement missed: an apparent false positive—a commuting zone the model flags but where no laboratory was seized—may be a true producer that evaded detection rather than an error, a point also raised in parallel work by Campedelli et al. (2024) for mafia infiltration. At the same time, I am cautious about pushing recall too far. Because the label is itself enforcement-based, maximizing recall against it

risks fitting the *enforcement* process—where the state happened to seize laboratories—rather than latent production, reproducing the very selection the exercise is meant to correct. Given this tension, rather than committing to a single optimal threshold I report estimates across a range of thresholds ($\tau \in \{0.2, 0.3, 0.4, 0.5\}$) and treatment definitions, and read the resulting spread as a robustness band.

Figure E.3: Recall–precision trade-off across classification thresholds



Notes: Recall (sensitivity) and precision of the random-forest predicted-exposure classifier as a function of the classification threshold τ, taking the baseline indicator (any methamphetamine laboratory seized before 2015) as the ground truth. Recall is the share of seizure commuting zones the classifier flags; precision is the share of flagged commuting zones that had a recorded seizure. Because seizures understate true production, precision is a lower bound—some apparent false positives may be undetected producers. Dashed lines mark the 0.25 threshold and the Youden-optimal threshold (0.19).
Source: Author's calculations.

↩ Return to text

E.6 Treatment Definitions

I construct two types of alternative treatment measures:

ML-Predicted Treatment. Classifies CZs based solely on predicted probabilities:

$$\hat{D}_i^{ML} = \mathbf{1}[\hat{P}(D_i^* = 1|X_i) > \tau] \tag{47}$$

where τ is the classification threshold; I report results for $\tau \in \{0.2, 0.3, 0.4, 0.5\}$.

Augmented Treatment. Combines baseline seizures with ML predictions:

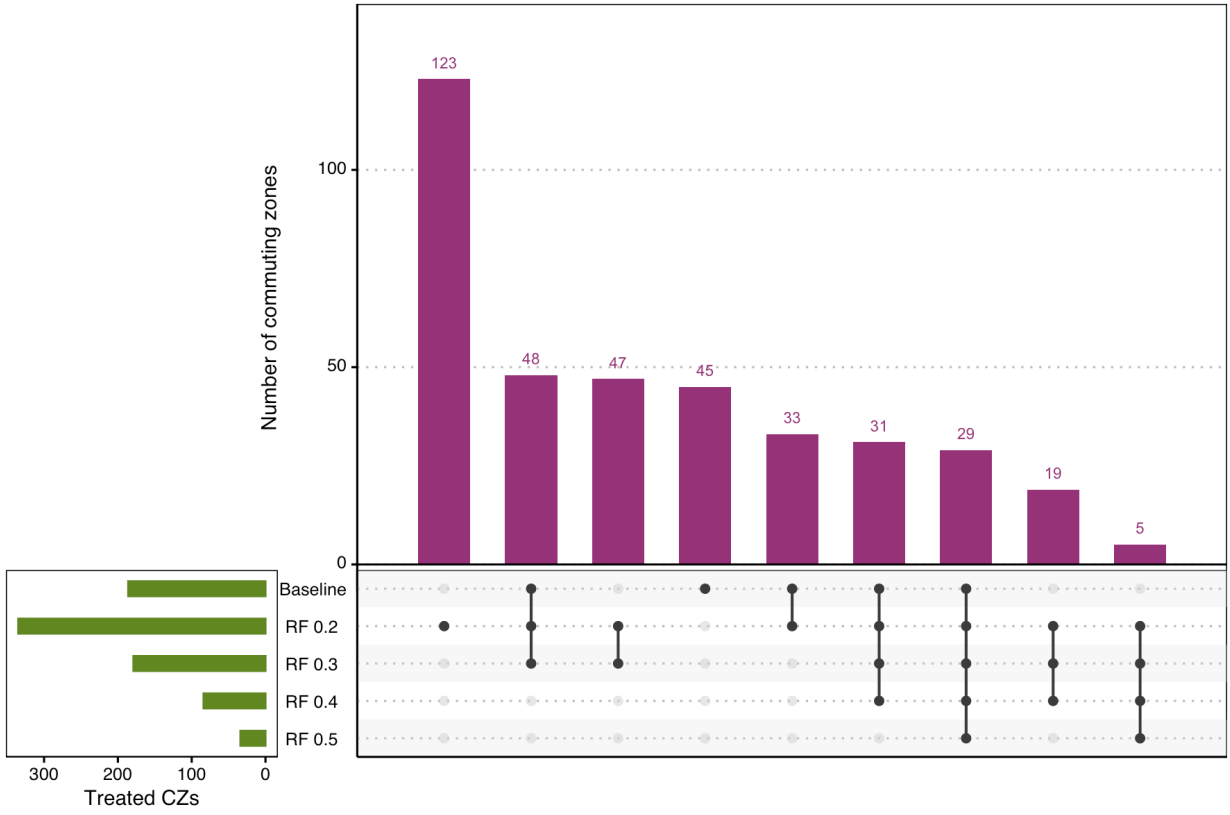
$$\hat{D}_i^{Aug} = \max\{D_i, \hat{D}_i^{ML}\} \tag{48}$$

This ensures 100% recall of baseline seizures while adding CZs identified as likely producers despite no observed seizures.

E.7 Results

Figure E.4 shows how the treatment definitions overlap across classification thresholds.

Figure E.4: Baseline and machine-learning treatment definitions overlap across thresholds



Notes: Upset plot of the overlap between the baseline seizure-based indicator (any methamphetamine laboratory seized before 2015) and the random-forest predicted-exposure indicator at thresholds 0.2, 0.3, 0.4, and 0.5. The left-hand bars give the size of each set; the top bars give the size of each intersection (the combination of sets marked by the dots below), i.e. the number of commuting zones falling in exactly that overlap. Higher thresholds yield smaller predicted sets that nest inside the lower-threshold sets.
Source: Author's calculations.

↩ Return to text

Overlap with Baseline. At the 0.25 threshold the random-forest classifier flags roughly 245 commuting zones (about 14% of the sample), recovering two-thirds of the baseline seizure zones (recall ≈ 0.67) while adding a comparable number of zones with no recorded seizure. As the threshold rises the predicted set shrinks and nests inside the lower-threshold sets, so both its size and its overlap with the baseline fall.

Augmented Measures. By construction, augmented measures retain every baseline seizure zone (100% recall of the baseline) and add the model-flagged zones; the augmented count therefore grows as the threshold falls.

F Classification of Tertiary Programs by Chemistry Intensity

To measure how synthetic drug production affects educational choices, I classify all tertiary programs in Mexico's ANUIES database (2015-2023) according to their chemistry intensity.

F.1 Step 1: Identifying Relevant Fields

I first restrict attention to programs in STEM and Health (STEM+H) fields using the International Standard Classification of Education (ISCED) 2013 framework. Specifically, I focus on programs in four broad fields to define STEM+H programs:

- Natural Sciences, Mathematics and Statistics
- Engineering, Manufacturing and Construction
- Agriculture, Forestry, Fisheries and Veterinary
- Health and Welfare

F.2 Step 2: Defining Reference Programs Using UNAM Curricula

To systematically assess chemistry intensity across thousands of unique program names, I use Universidad Nacional Autónoma de México (UNAM) as a reference institution. UNAM is Mexico's largest and most prestigious public university, offering a comprehensive range of STEM+H programs with publicly available curricula, (*mapas curriculares*).

For all programs in STEM+H, I:

1. Identified all UNAM programs in the ANUIES data in the STEM+H ISCED classifcations
2. Downloaded the official curriculum (*mapa curricular*) for each program from their website
3. Manually reviewed each curriculum to identify chemistry courses, recording:
    - Total chemistry courses throughout the program
    - Chemistry courses in the first year
    - Chemistry or other courses particularly useful for synthetic drug manufacturing

This process yielded 116 reference UNAM programs with detailed chemistry course information.

F.3 Step 3: Mapping All Programs to UNAM Reference Programs

I then mapped each of the 1,308 unique program names in the ANUIES data to one of the UNAM reference programs (or to no UNAM program if outside STEM+H fields).

F.3.1 Text Preprocessing

To standardize program names across institutions and enable systematic matching, I applied the following preprocessing pipeline to all program names:

1. **Degree term removal**: Removed common degree designations ("licenciatura", "maestría", "técnico superior universitario", etc.) that do not convey substantive program content
2. **Normalization**:
    - Converted all text to lowercase using Spanish locale
    - Removed accents and diacritical marks (e.g., "á" → "a")
    - Removed all non-alphanumeric characters except spaces
3. **Stopword removal**: Removed common Spanish stopwords (articles, prepositions, conjunctions) using the `stopwords` package Spanish dictionary
4. **Suffix stripping**: Applied a rule-based stemming algorithm to remove common Spanish suffixes (e.g., "-aciones", "-imientos", "-ista", "-ología") to group morphological variants. Suffixes were removed in order from longest to shortest to avoid partial matches. For example:
    - "ingeniería" → "ingenier"
    - "química" → "quim"
    - "biología" → "biolog"
5. **Component extraction**: Extracted the first three substantive words from each stemmed program name as `program_name_rw1`, `program_name_rw2`, and `program_name_rw3`

This preprocessing pipeline reduces variation in program naming conventions across institutions while preserving substantive content. For example, "Licenciatura en Ingeniería Química Industrial" becomes the stemmed components: `ingenier`, `quim`, `industri`.

F.3.2 Manual Rule-Based Matching

Using the preprocessed program name components and ISCED field classifications, I manually created detailed mapping rules to assign each program to the most appropriate UNAM reference program. This process was largely manual rather than algorithmic, requiring careful review of program name patterns and domain knowledge of Mexican higher education.

The mapping rules use conditional logic based on combinations of the three word stems and ISCED field classifications. For each UNAM reference program, I defined multiple matching patterns to capture the various ways institutions name similar programs. The rules account for:

- **Variant naming conventions**: Different institutions use different terminology for the same program (e.g., "ingeniería química", "químico industrial", "ingeniería en procesos químicos")
- **Specializations within fields**: Some programs require matching all three word stems to distinguish specializations (e.g., distinguishing "ingeniería química" from "ingeniería química de alimentos")
- **Field-dependent disambiguation**: The same first word can map to different programs depending on ISCED field (e.g., "biotecnología" maps to "ciencias agrogenómicas" in Agriculture/Veterinary fields but to "ciencias genómicas" in Natural Sciences)
- **Compound matching patterns**: Some programs are identified through specific two- or three-word combinations or through pattern matching in the cleaned program name

Example matching rules for selected programs:

- **Ingeniería química**: Matches programs with (`program_name_rw1 == "ingenier"` AND `program_name_rw2 == "quim"`), OR (`program_name_rw1 == "quimic"` AND `ISCED_field_name == "Engineering..."`)
- **Ingeniería química de alimentos**: Matches programs with (`program_name_rw1 == "quim"` AND `program_name_rw2 == "aliment"`), OR (`program_name_rw1 == "ingenier"` AND `program_name_rw2 == "quim"` AND `program_name_rw3 == "aliment"`)
- **Biología**: Matches programs with `program_name_rw1` in {"biolog", "hidrobiolog", "microbiolog"}, OR (`program_name_rw1 == "ci"` AND `program_name_rw2 == "biolog"`)

- **Ciencias agrogenómicas vs. genómicas**: Programs with `program_name_rw1 == "biotecnolog"` map to "ciencias agrogenómicas" if in Agriculture/Veterinary fields, otherwise to "ciencias genómicas"

This mapping required creating multiple matching rules across UNAM reference programs. The process was iterative: I reviewed unmatched programs, identified common patterns, and added new rules to improve coverage. Programs that could not be confidently matched to any UNAM reference remain unmapped and receive zero values for all chemistry intensity measures.

The mapping prioritizes precision over recall—I only assign a match when there is clear correspondence in subject matter. This conservative approach minimizes misclassification but means some STEM+H programs may be incorrectly classified as having zero chemistry intensity if they could not be matched to a UNAM reference.

F.4 Step 4: Constructing Chemistry Intensity Measures

For each program, I construct three measures of chemistry intensity:

1. `n_chem_courses_tot`: Total number of chemistry courses throughout the program
2. `n_chem_courses_y1`: Number of chemistry courses in the first year
3. `n_drug_courses_tot`: Number of chemistry courses particularly useful for synthetic drug manufacturing

Programs that could not be mapped to any UNAM reference receive values of zero for these measures.

F.5 Step 5: Classification into STEM+H Chemistry vs. Non-Chemistry

For the main analysis, I create a binary classification:

- **STEM+H with chemistry**: Programs with `n_chem_courses_tot` > 0
- **STEM+H without chemistry**: Programs in STEM+H fields with `n_chem_courses_tot` $== 0$
- **Non-STEM+H**: All other programs

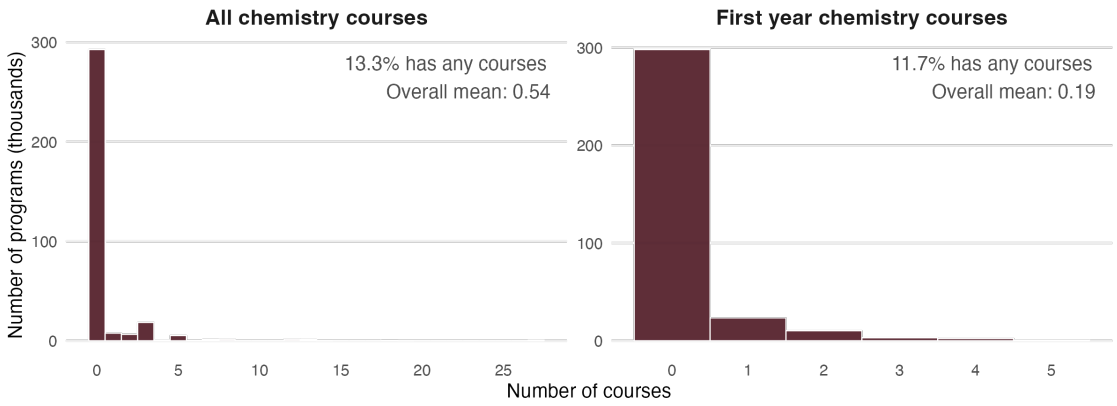
F.6 Distribution of Chemistry Intensity

Figure F.1 shows the distribution of chemistry intensity across programs. Among all fields, 13.3% of programs have at least one chemistry course. Looking at only STEM+H programs, almost half of the programs have at least one chemistry course, the mean number of chemistry courses is 1.96, with substantial variation ranging from 0 to 27 courses. The distribution is highly right-skewed, with most STEM+H programs requiring either zero or very few chemistry courses, while a smaller subset of programs require intensive chemistry training. First-year chemistry requirements follow a similar pattern but are more concentrated, with 42.5% of STEM+H programs requiring at least one chemistry course in the first year and a mean of 0.7 first-year chemistry courses. This concentration of chemistry requirements in the first year suggests that early exposure to chemistry knowledge is a key distinguishing feature of any chemistry-related programs.

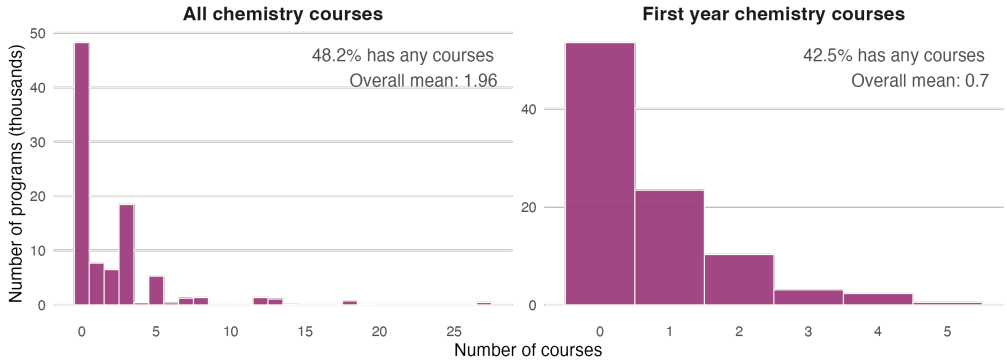
Figure F.2 shows the 15 most common tertiary programs, separated by chemistry content classification. Among the most popular programs, Non-STEM+H fields dominate, with business administration, law, and psychology among the top programs. The most common STEM+H programs with chemistry content are industrial engineering, nursing, and nutrition, while the most common STEM+H programs without chemistry include architecture, mechatronics engineering, and civil engineering.

Figure F.1: Chemistry intensity of tertiary education programs

(a) All fields



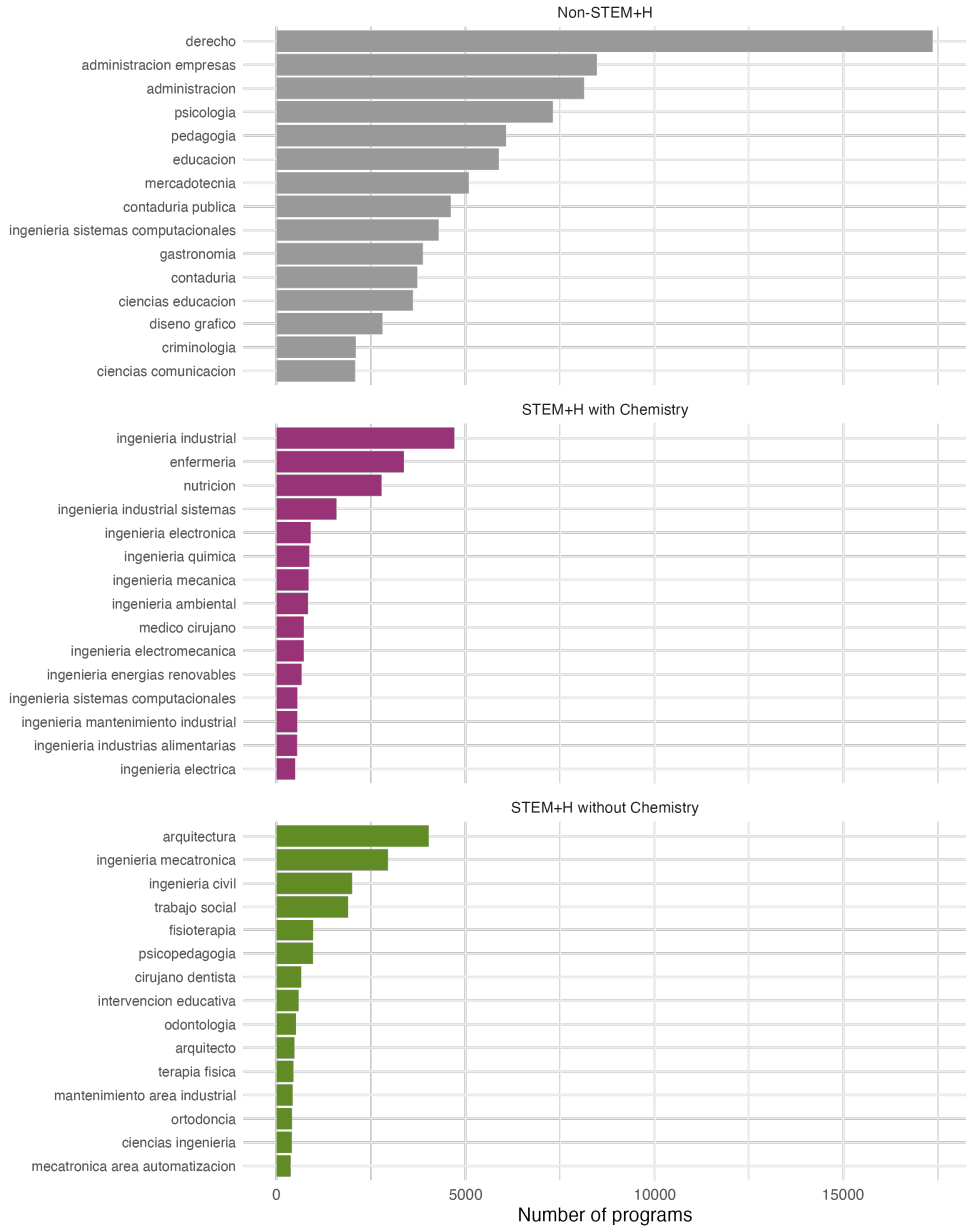
(b) STEM+H programs



Notes: The figure shows the distribution of chemistry course requirements across tertiary education programs in Mexico, 2015-2023. Panel (a) includes all programs across all fields; Panel (b) restricts to programs in STEM+H fields (Natural Sciences, Mathematics and Statistics; Engineering, Manufacturing and Construction; Agriculture, Forestry, Fisheries and Veterinary; and Health and Welfare). Chemistry intensity is measured by counting chemistry courses in program curricula from Universidad Nacional Autónoma de México (UNAM), which are used as reference curricula for mapping all programs in the ANUIES database. Programs that could not be mapped to UNAM reference programs are assigned zero chemistry courses. The left panels show total chemistry courses throughout the program; the right panels show chemistry courses required in the first year only. The share of programs with at least one chemistry course and the average number of chemistry courses are shown in the top right of each panel.
Source: Author's calculations based on ANUIES tertiary enrollment data (2015-2023) and UNAM program curricula.

↩ Return to text

Figure F.2: Most common tertiary education programs my STEM+H and chemistry content



Notes: The figure shows the 15 most common tertiary education programs in Mexico (2015-2023) by field classification. Programs are classified into three categories: Non-STEM+H (gray), STEM+H with chemistry (purple), and STEM+H without chemistry (green). STEM+H fields include Natural Sciences, Mathematics and Statistics; Engineering, Manufacturing and Construction; Agriculture, Forestry, Fisheries and Veterinary; and Health and Welfare. Chemistry content is determined by mapping programs to UNAM reference curricula and counting chemistry course requirements. Programs with at least one chemistry course are classified as "STEM+H with chemistry." The sample includes all program-institution-year observations in the ANUIES database.
Source: Author's calculations based on ANUIES tertiary enrollment data (2015-2023) and UNAM program curricula.

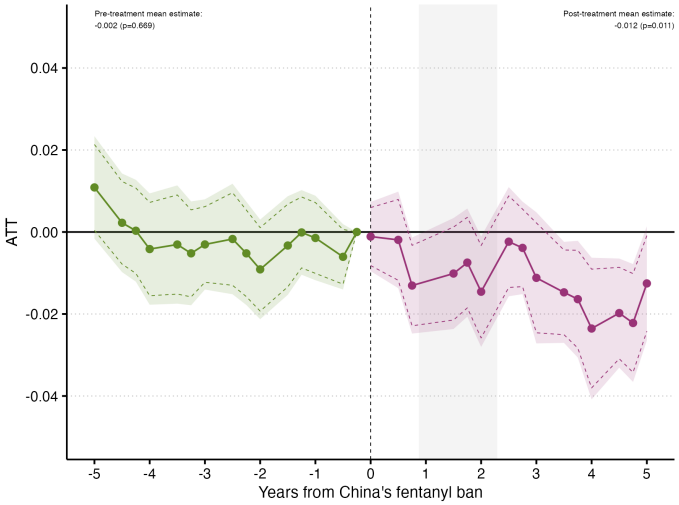
↩ Return to text

G Additional Results

G.1 Middle school dropout

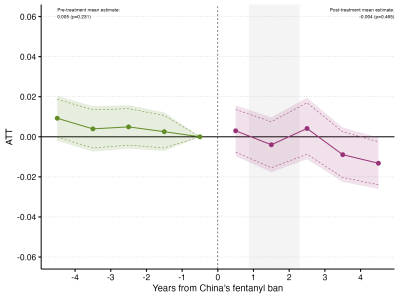
Figure G.1: Event Study of Exposure to Synthetic Drug Production on Middle School Dropout

(a) All quarters

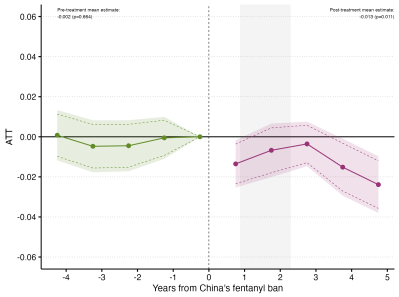


(b) By quarter

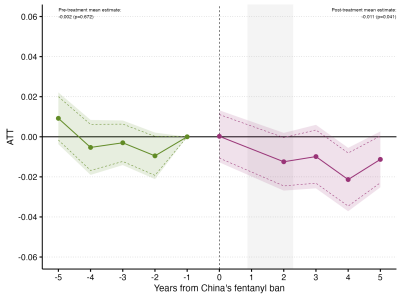
Q4 (Oct–Dec)



Q1 (Jan–Mar)



Q2 (Apr–Jun)



Notes: Event-study estimates of the effect of exposure on school dropout at middle school age (12–14). The specifications match Figure 3: panel (a) pools all quarters with commuting zone, year-quarter, and quarter fixed effects; panel (b) is estimated separately for each calendar quarter with commuting zone and year fixed effects. All specifications control for gender, age, and household education and use survey weights. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level. The shaded region marks the COVID-19 school-closure period.
Source: Author's calculations based on ENOE.

↩ Return to text

Table G.1: Effect of Exposure to Synthetic Drug Production on Middle School Dropout

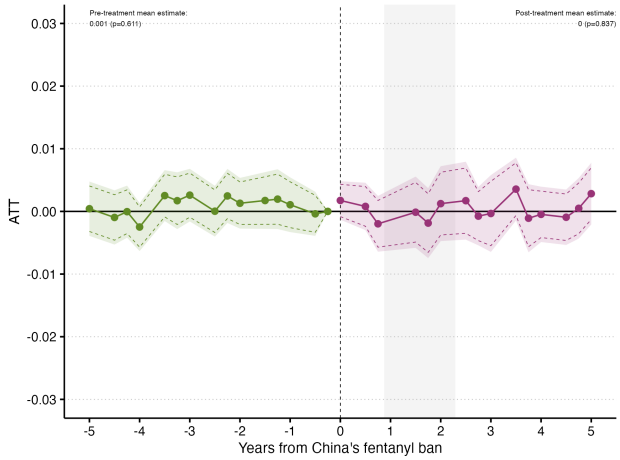
	All quarters				By quarter		
	(1)	(2)	(3)	(4)	Oct-Dec	Jan-Mar	Apr-Jun
Exposure × Post	-0.012*	-0.012*	-0.012**	-0.012*	-0.010*	-0.014**	-0.011*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Exposure × Post × Female				0.000			
				(0.004)			
Female		-0.009***	-0.009***	-0.005*	-0.007***	-0.008***	-0.012***
		(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Age		0.021***	0.020***	0.020***	0.019***	0.021***	0.019***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Household avg. education			-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	665,100	665,100	665,100	665,100	214,645	230,717	219,738
Mean of dep. var.	0.057	0.057	0.057	0.057	0.051	0.057	0.063

Notes: The dependent variable is an indicator for not being enrolled in school among individuals of middle school age (12–14). Columns follow Table 2: columns (1)–(4) pool all quarters with commuting zone, year-quarter, and quarter fixed effects; columns (5)–(7) are estimated separately for each calendar quarter with commuting zone and year fixed effects and full controls. Observations are weighted by survey weights. Mean of dep. var. is the mean in control commuting zones before the ban. Standard errors clustered at the commuting zone level are reported in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

↩ Return to text

G.2 Placebo: primary school age

Figure G.2: Event studies of the Synthetic Drug Expansion on Primary Age School Dropout

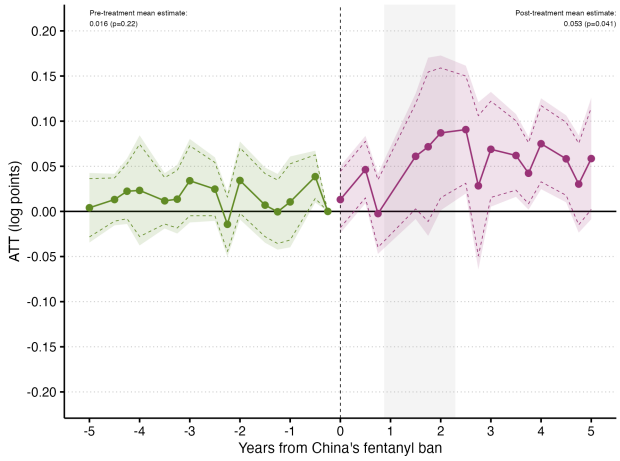


Notes: Placebo event study for primary school age children (6–11), who are not at the margin of leaving school for work. The dependent variable is an indicator for not being enrolled in school. The specification matches Figure 3: commuting zone, year-quarter, and quarter fixed effects, controls for gender, age, and household education, survey weights, and the period before the fentanyl ban as the reference. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level.
Source: Author's calculations based on ENOE.

↩ Return to text

G.3 Time use

Figure G.3: Event study of exposure to synthetic drug production on weekly time spent studying



Notes: Event-study estimates of the effect of exposure on weekly time spent studying among school-age youth (12–17), from a Poisson (PPML) regression; coefficients are proportional effects (log points). Exposure is an indicator for any methamphetamine laboratory seized in the commuting zone before 2015, interacted with period indicators relative to the period before the fentanyl ban. The model includes commuting zone, year-quarter, and quarter fixed effects and controls for gender, age, and household education; observations are weighted by survey weights. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level. The shaded region marks the COVID-19 school-closure period.
Source: Author's calculations based on the ENOE time-use module.

↩ Return to text

Table G.2: Impact of Exposure to Synthetic Drug Production on Time Use

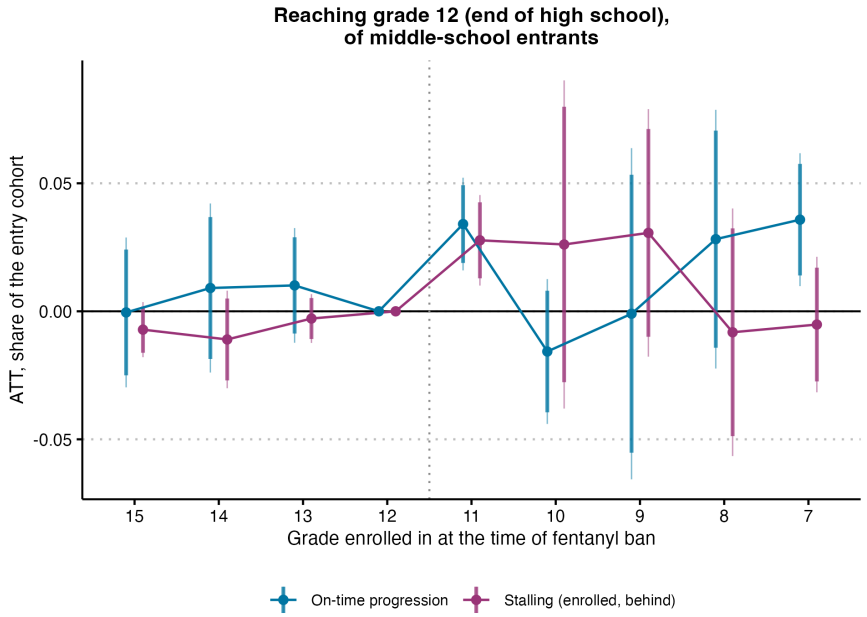
Statistic	Study (Any)	Study (Min)	Care (Min)	Shopping (Min)	Chores (Min)	Volunteer (Min)
Middle School						
Exposure × Post	0.0154*	0.0298*	-0.0153	-0.1235	-0.0528**	-0.0434
	(0.0080)	(0.0177)	(0.0516)	(0.0807)	(0.0259)	(0.1681)
Model	LPM	Poisson	Poisson	Poisson	Poisson	Poisson
Control Mean	0.8586	1864.6082	46.1162	29.7047	361.2506	1.3594
N	655,771	654,047	651,645	654,418	653,028	573,318
High School						
Exposure × Post	0.0271***	0.0481**	-0.0057	-0.0676	-0.0452***	0.2445
	(0.0080)	(0.0190)	(0.0550)	(0.0580)	(0.0173)	(0.1821)
Model	LPM	Poisson	Poisson	Poisson	Poisson	Poisson
Control Mean	0.6582	1483.7514	95.0637	40.3341	480.0166	2.8078
N	674,457	672,107	670,775	673,212	670,881	614,398

Notes: Difference-in-differences estimates of the effect of exposure to synthetic drug production on weekly time use among school-age youth (12–17 years). Study (Any) is an indicator for reporting any time spent studying and is estimated by a linear probability model; the remaining (minutes) outcomes are estimated by Poisson, so the reported coefficients are proportional effects (semi-elasticities). The control mean is reported in minutes for reference. Exposure indicates any methamphetamine laboratory seized in the commuting zone before 2015; Post indicates years after 2019. All models include commuting zone, year-quarter, and quarter fixed effects and control for gender, age, and household education; observations are weighted by survey weights. Standard errors clustered at the commuting zone level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

↩ Return to text

G.4 Grade progression

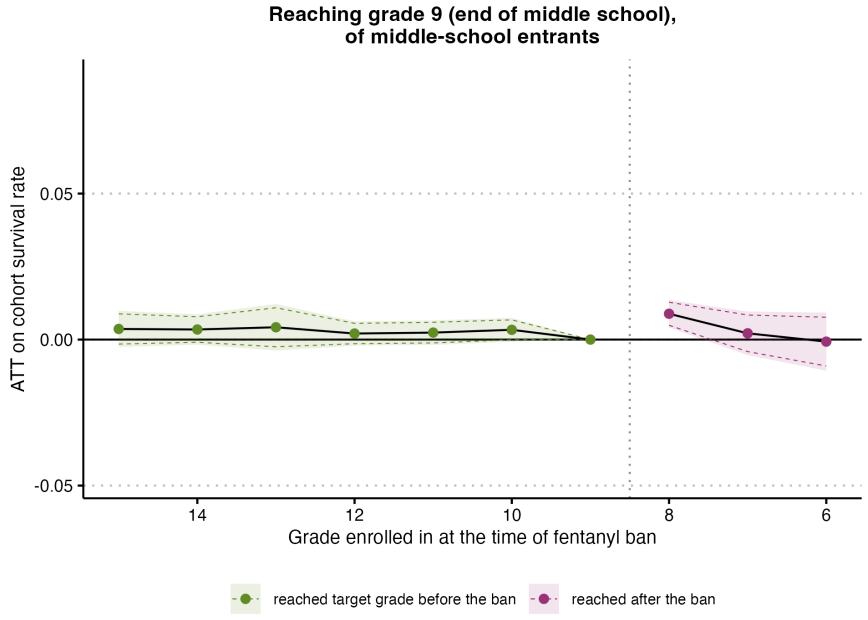
Figure G.4: Reaching grade 12: on-time progression and stalling



Notes: Cohort event study of Figure 4, splitting the share of the middle-school entry cohort still enrolled at grade 12 into students who progressed on time and students who are still enrolled but behind (stalling): $S_{ret} = S_{prog} + S_{rep}$. Cohorts are indexed by the grade they were enrolled in at the ban; the reference is the cohort in grade 12 at the ban. Thick and thin bars are 90% and 95% confidence intervals from standard errors clustered at the commuting zone level. *Source:* Author's calculations based on Formato 911.

↩ Return to text

Figure G.5: Reaching the end of middle school (grade 9)



Notes: Cohort event study of the share of the middle-school entry cohort (grade 7) that remains continuously enrolled through grade 9. Cohorts are indexed by the grade they were enrolled in at the ban; the reference is the last cohort to reach grade 9 before the ban. Commuting-zone and grade-at-ban fixed effects, cohorts weighted by size. Bands are 95% confidence intervals and dashed lines 90% confidence intervals, from standard errors clustered at the commuting zone level.
Source: Author's calculations based on Formato 911.

↩ Return to text

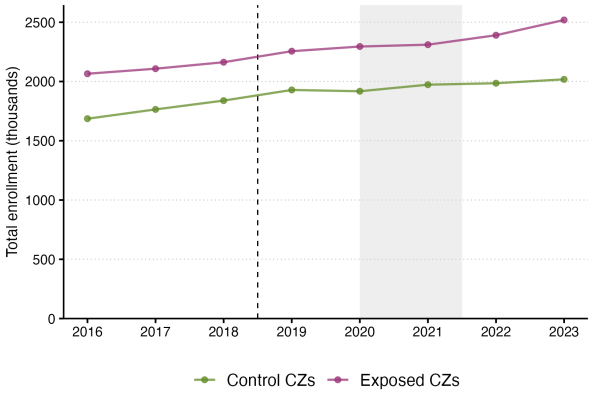
G.5 Raw trends in administrative data

Figure G.6: Raw trends in administrative education data, exposed and control commuting zones

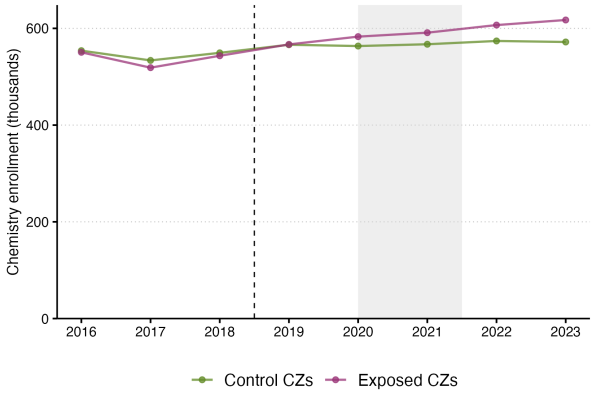
(a) Share of the entry cohort reaching grade 12



(b) Total university enrollment



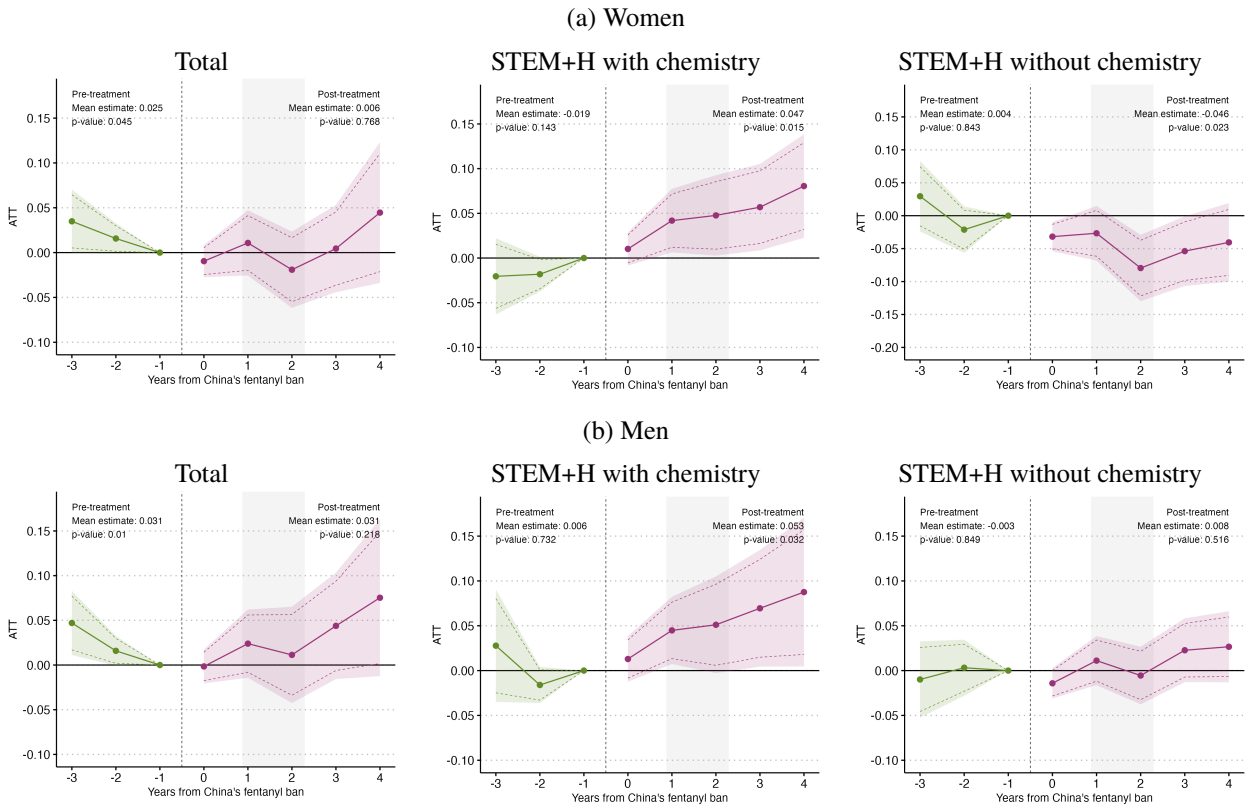
(c) Enrollment in chemistry-intensive programs



Notes: Means or totals by group and year, without controls. Panel (a) is the share of each middle-school entry cohort that reaches grade 12, by the year the cohort reaches it, weighted by cohort size (Formato 911). Panels (b) and (c) are ANUIES enrollment assigned to origin commuting zones with the 2010 migration matrix, summed by group, in thousands of students. Exposed commuting zones had a methamphetamine laboratory seized before 2015. The dashed line marks the fentanyl ban and the shaded area the COVID-19 school closures.
Source: Author's calculations based on Formato 911, ANUIES, and the 2010 Population Census.

G.6 University enrollment

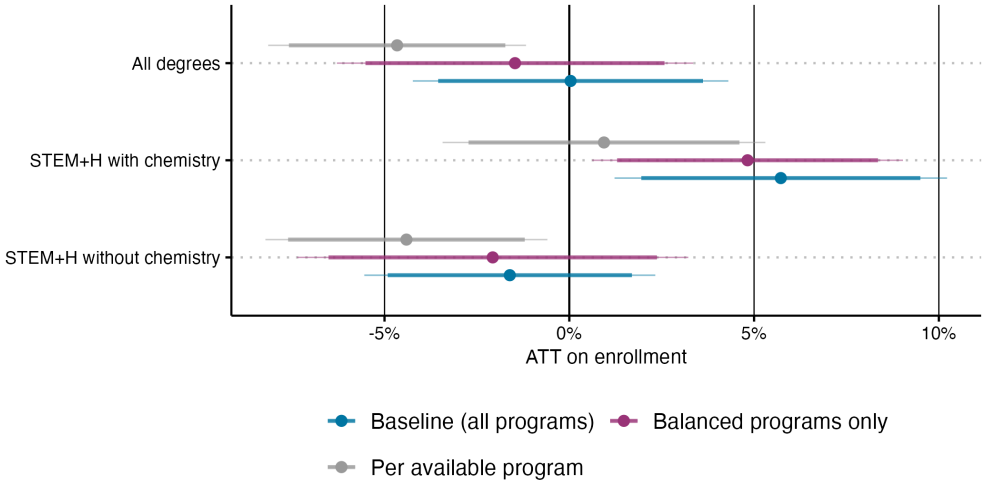
Figure G.7: Impact of the Synthetic Drug Expansion on Tertiary Enrollment, by Gender



Notes: Event-study estimates as in Figure 5, separately for women and men. Enrollment is assigned to origin commuting zones with the 2010 migration matrix. Poisson (PPML) with commuting zone and year fixed effects, reference year 2018; coefficients are proportional effects. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level.
Source: Author's calculations based on ANUIES tertiary enrollment data.

↩ Return to text

Figure G.8: Tertiary enrollment and the supply of programs



Notes: Poisson (PPML) difference-in-differences estimates of the effect of exposure on tertiary enrollment by field. "Baseline" uses all programs. "Balanced programs only" restricts to programs observed in every year. "Per available program" adds the log number of programs in the field, assigned to origins with the same migration matrix as enrollment, as an offset. All specifications include commuting zone and year fixed effects. Thick and thin bars are 90% and 95% confidence intervals from standard errors clustered at the commuting zone level.
Source: Author's calculations based on ANUIES tertiary enrollment data.

↩ Return to text

G.7 Labor market

Table G.3: Effects on the composition of the workforce: education and geographic origin

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Educational composition of the workforce (average years of schooling)						
	Dependent variable: average years of schooling					
		By occupation group				
					STEM by chemistry content	
	All	Professional	Non-STEM	STEM	No	Yes
Exposure × Post	-0.0365	0.0146	0.0224	-0.1518***	-0.0843	-0.2849**
	(0.0399)	(0.0374)	(0.0395)	(0.0586)	(0.0520)	(0.1370)
Observations (CZ × year)	9,657	8,196	8,138	3,729	3,328	1,868
Mean (control)	10.879	14.814	14.666	15.902	16.217	14.361
No. of commuting zones	1,235	1,162	1,159	740	694	443
Panel B. Geographic origin of the workforce (out-of-state share)						
	Dependent variable: out-of-state (migrant) share					
		By occupation group				
					STEM by chemistry content	
	All	Professional	Non-STEM	STEM	No	Yes
Exposure × Post	-0.0089*	-0.0000	-0.0001	-0.0023	-0.0023	-0.0021
	(0.0047)	(0.0057)	(0.0053)	(0.0124)	(0.0131)	(0.0175)
Observations (CZ × year)	9,657	8,199	8,141	3,730	3,328	1,870
Mean (control)	0.2091	0.2153	0.2100	0.2584	0.2638	0.2321
No. of commuting zones	1,235	1,162	1,159	740	694	443

Notes: Difference-in-differences estimates; standard errors clustered by commuting zone in parentheses. Both panels are commuting-zone × year weighted OLS (weights = workforce size) with commuting-zone and year fixed effects, estimated on prime-age (25–54) employed workers. *Exposure* equals one for commuting zones with a methamphetamine laboratory seized before 2015; *Post* equals one from 2019 (China's fentanyl ban). Each outcome is computed over the workforce of the occupation group heading the column: *Panel A* is the average years of schooling of that group's workers (the coefficient is a change in years), and *Panel B* is the share of that group's workers born in a state other than their state of residence (a change in percentage points). Columns follow the occupation grouping of Table 4, Panel B: (1) all employed, (2) professional occupations (ISCO-08 major group 2), (3) non-STEM professional, (4) STEM, and (5)–(6) STEM split by chemistry content, with (6) the chemistry occupations (chemists, chemical engineers, and chemical-plant operators). The number of commuting zones falls across columns because thinner occupation groups are present in fewer zones. $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$.

↩ Return to text

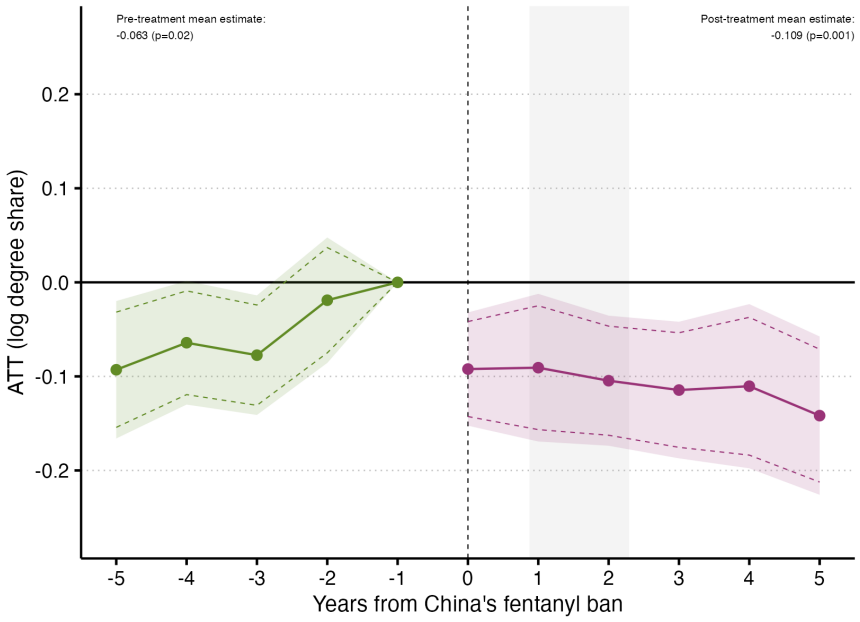
Table G.4: Effects on the stock of degree holders, aged 30–54 at the time of the shock

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: share of the cohort holding a degree				
		By field of study			
				STEM by chemistry content	
	Any tertiary	Non-STEM	STEM	No	Yes
Exposure × Post	0.0020	0.0067	-0.0366*	-0.0253	-0.0461**
	(0.0149)	(0.0164)	(0.0193)	(0.0280)	(0.0226)
Observations (CZ × year)	9,236	9,260	7,998	6,873	7,282
Control share (degree)	0.2243	0.2211	0.0548	0.0207	0.0343
No. of commuting zones	1,093	1,097	838	671	738

Notes: Poisson (PPML) difference-in-differences estimates of the share of a fixed birth cohort holding a degree; standard errors clustered by commuting zone in parentheses. The cohort is people aged 30–54 *in 2019*, followed across 2012–2024 by age in 2019, so their attainment is predetermined and a decline cannot reflect new schooling—only mobility or observability. Estimated on commuting-zone × year cells with an offset for log cohort size, commuting-zone and year fixed effects; *unconditional* on employment. *Exposure* equals one for commuting zones with a methamphetamine laboratory seized before 2015; *Post* equals one from 2019. Column (1) is any tertiary degree; columns (2)–(5) are by field of study, with (4)–(5) splitting STEM by chemistry content ((5) = chemistry-course fields, including pharmacy). Any tertiary (1) and non-STEM (2) are placebos: a decline concentrated in STEM and chemistry, with the placebos flat, rules out generic educated out-migration or survey non-response, which would fall on all fields alike. $*$ p < 0.1, $**$ p < 0.05, $***$ p < 0.01.

↩ Return to text

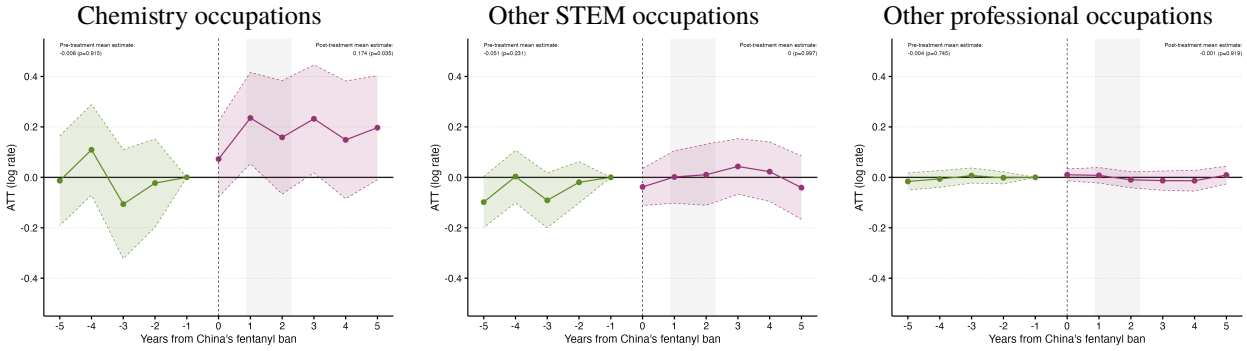
Figure G.9: Stock of degree holders: alternative chemistry grouping



Notes: Event study as in Figure 9b, for the share of the fixed 2019 cohort (aged 30–54 in 2019) holding a degree in any field with chemistry courses, including pharmacy. Poisson pseudo-maximum-likelihood at the commuting-zone × year level with an offset for log cohort size, commuting-zone and year fixed effects. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level.
Source: Author's calculations based on ENOE.

↩ Return to text

Figure G.10: Employment event studies by occupation group



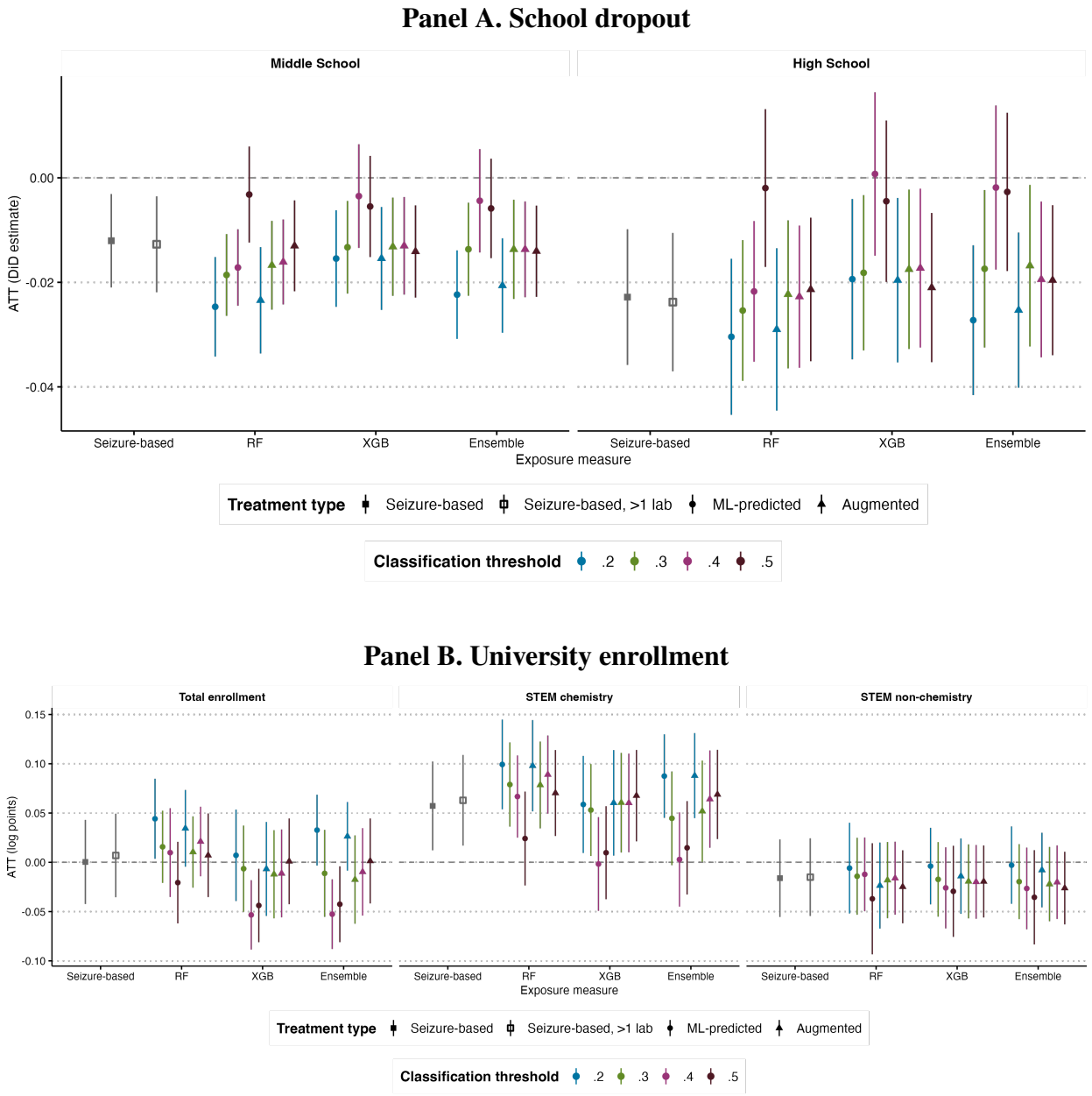
Notes: Event studies of the number of prime-age (25–54) workers employed in each occupation group, estimated by Poisson pseudo-maximum-likelihood at the commuting-zone × year level with an offset for log employment, so coefficients are proportional changes in the occupation's share of employment. Commuting-zone and year fixed effects; reference year $t = -1$. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level.
Source: Author's calculations based on ENOE.

↩ Return to text

H Robustness checks

H.1 Alternative treatment definitions

Figure H.1: DiD estimates of alternative treatment measures, all definitions

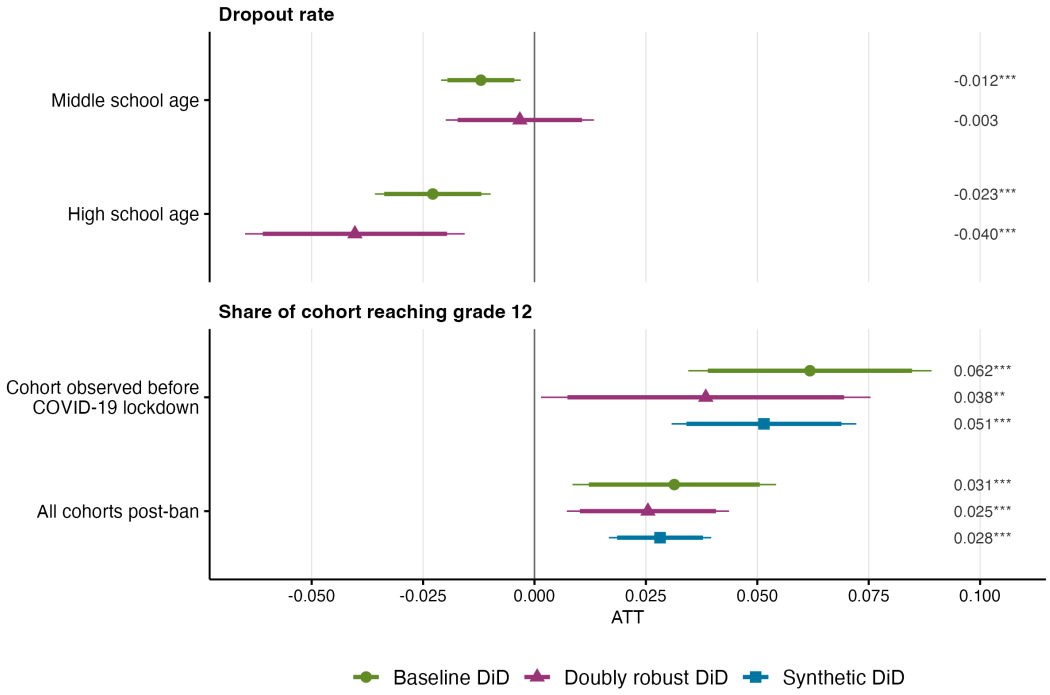


Notes: Difference-in-differences estimates of the effect of exposure on school dropout (Panel A) and tertiary enrollment (Panel B) under all alternative definitions of treatment. Within each panel, the x-axis indicates the exposure measure: the baseline seizure-based indicator (any methamphetamine laboratory seized before 2015), and machine-learning predictions of latent production from a random forest (RF), gradient boosting (XGB), and their ensemble. Point shape denotes the treatment type—seizure-based, ML-predicted (a commuting zone is treated if its predicted probability exceeds the threshold), or augmented (the maximum of the seizure-based and ML-predicted indicators); color denotes the classification threshold (.2, .3, .4, .5). Panel A reports linear-probability DiD coefficients for middle and high school dropout (commuting zone, year-quarter, and quarter fixed effects; controls for gender, age, and household education). Panel B reports Poisson (PPML) coefficients—proportional effects—for total enrollment, chemistry-intensive STEM+H, and STEM+H without chemistry (commuting zone and year fixed effects). Bars are 95% confidence intervals from standard errors clustered at the commuting zone level.

↩ Return to text

H.2 Alternative estimators

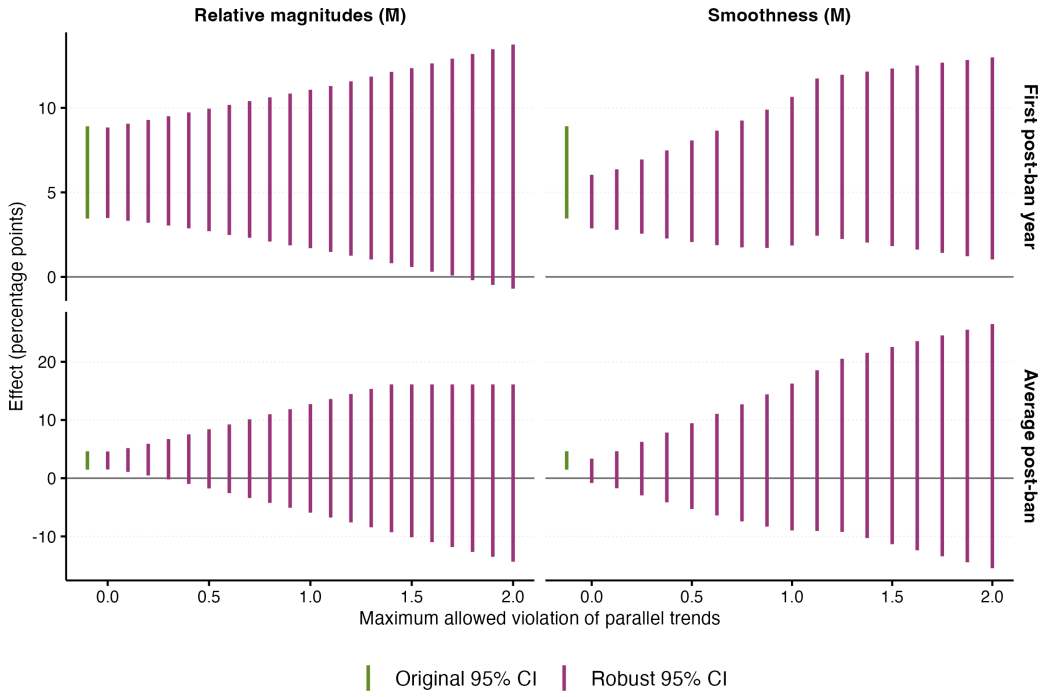
Figure H.2: Baseline, doubly robust, and synthetic difference-in-differences estimates



Notes: Baseline difference-in-differences estimates compared with the doubly robust estimator of Sant'Anna and Zhao (2020), which reweights control commuting zones on observable baseline characteristics, and, for the share of the cohort reaching grade 12, with synthetic difference-in-differences. Upper panel: ENOE dropout at middle and high school age. Lower panel: share of the middle-school entry cohort reaching grade 12, for the cohort that reached it before the COVID-19 closures and for all post-ban cohorts. Thick and thin bars are 90% and 95% confidence intervals.
Source: Author's calculations based on ENOE and Formato 911.

↩ Return to text

Figure H.3: Sensitivity to violations of parallel trends: reaching grade 12



Notes: Robust 95% confidence sets of Rambachan and Roth (2023) for the effect of exposure on the share of the middle-school entry cohort reaching grade 12, for the first post-ban cohort (top) and the average over post-ban cohorts (bottom). The left column bounds post-ban violations of parallel trends relative to the largest pre-ban violation (relative magnitudes, $\bar{M}$); the right column bounds changes in the slope of the trend (smoothness, M). Green bars are the original 95% confidence intervals.
Source: Author's calculations based on Formato 911.

↩ Return to text

H.3 Minimum wage increases

Mexico raised the minimum wage substantially from 2019, from 88.36 pesos per day in 2018 to 102.68 pesos in 2019 and 248.93 pesos by 2024. In the municipalities of the northern border zone (Zona Libre de la Frontera Norte, ZLFN), the minimum wage doubled to 176.72 pesos in 2019 (Calderón et al., 2023). Figure H.4 shows its evolution and the municipalities in the ZLFN. Higher minimum wages can raise dropout by making work more attractive for young people, or reduce it by raising household income.

Exposed commuting zones were slightly less affected by the increases. Table H.1 regresses an indicator for workers whose 2018 wage was below the new minimum on exposure. Workers in exposed commuting zones were 0.4 percentage points less likely to be affected by the 2019 increase and 0.8 percentage points less likely to be affected by the 2021 level, and the differences are smaller

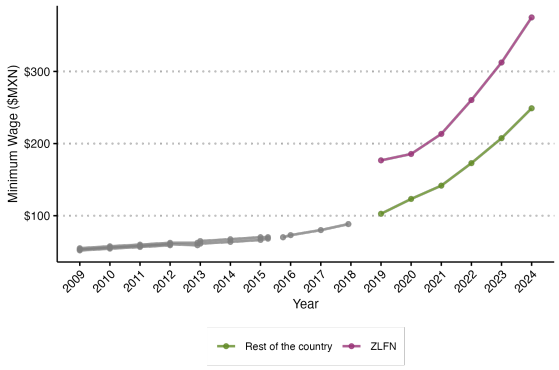
outside the ZLFN (0.1 and 0.3 percentage points).[28] Since exposed areas were less affected, the minimum wage could explain the results only if it raised dropout in control areas.

In control areas, the minimum wage increases were associated with lower dropout. Table H.2 uses only control commuting zones and relates dropout, youth employment, and household income to the share of workers affected by the increases. A commuting zone where 10% more workers were affected by the 2021 level saw dropout fall by 1.7 percentage points at middle school age and 2.0 at high school age after 2019. The estimates for the 2019 level are also negative but not significant. Household labor income rose with exposure to the 2019 increase. This suggests that the income effect dominated, so any bias from the minimum wage works against finding lower dropout in exposed areas.

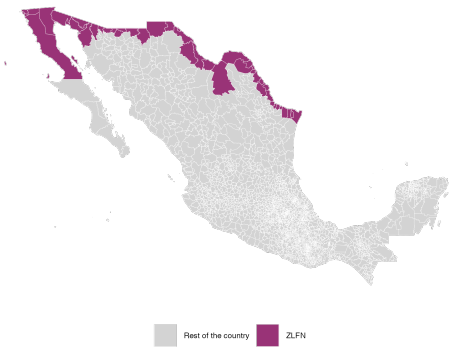
The effect of exposure does not change when I account for the minimum wage. Table H.3 reports the effect of exposure outside the ZLFN, where the increases were smaller, and when controlling for household labor income. Outside the ZLFN, exposure reduces dropout by 1.3 percentage points at middle school age and 2.3 at high school age, the same as in the full sample. Controlling for household income gives 1.1 and 2.0 percentage points. Within the ZLFN the estimates are smaller and imprecise, because only a few commuting zones are in it.

Figure H.4: Minimum wage increases

(a) Evolution of nominal minimum wage



(b) Minimum wage zones



Notes: Panel A shows the evolution of the minimum wage in current Mexican pesos. Starting in 2019, there are two different minimum wages, one for Zona Libre de la Frontera Norte (ZLFN), shown in purple, and another for the rest of the country, in green. Panel B shows a map of Mexican municipalities by minimum wage zone.
Source: Comisión Nacional de los Salarios Mínimos.

⟷ Return to text

[28] Haanwinckel (2025) evaluates several approaches to measuring regional minimum wage bite and finds that the fraction affected design is preferable to alternatives such as the effective minimum wage or gap measures, particularly when pre-treatment data is available to test for parallel trends.

Table H.1: Exposure to synthetic drug production and share of workers affected by minimum wage increases

	Dep. variable: Worker is affected by MW increase			
	Full sample		Excluding ZLFN	
	2019 MW	2021 MW	2019 MW	2021 MW
	(1)	(2)	(3)	(4)
Exposure	-0.004***	-0.008***	-0.001***	-0.003***
	(0.000)	(0.001)	(0.000)	(0.001)
Num.Obs.	705297	705297	664183	664183

Notes: This table reports OLS estimates of the correlation between exposure to clandestine drug labs and the share of workers affected by minimum wage increases. The dependent variable is a dummy equal to one if the worker's job is formal and their daily wage in 2018 was between the current minimum wage that would take effect in 2019 (columns 1 and 3) or 2021 (columns 2 and 4). Columns 1-2 use the full sample; columns 3-4 exclude municipalities in the Zona Libre de la Frontera Norte (ZLFN), which experienced larger minimum wage increases. Standard errors in parentheses. Observations are weighted using survey expansion factors. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.
Data: ENOE 2018.

⟷ Return to text

Table H.2: Minimum wage increases and dropout in the control group

	Dependent variable:							
	School Dropout		Works		HH Labor Income			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Middle school age								
FA by 2019 MW × Post	-0.074		-0.044		0.404+		0.547*	
	(0.111)		(0.097)		(0.218)		(0.262)	
FA by 2021 MW × Post		-0.168*		-0.092*		-0.350		-0.049
		(0.065)		(0.047)		(0.246)		(0.182)
Num.Obs.	394893	394893	394895	394895	322977	322977	310572	310572
Panel B: High school age								
FA by 2019 MW × Post	-0.123		-0.070		0.577+		0.934***	
	(0.158)		(0.319)		(0.348)		(0.267)	
FA by 2021 MW × Post		-0.200*		-0.148		-0.383		-0.093
		(0.088)		(0.104)		(0.257)		(0.209)
Num.Obs.	397663	397663	397670	397670	321500	321500	297019	297019

Notes: This table reports difference-in-differences estimates of the effect of exposure to minimum wage increases on school dropout and household labor income. The dependent variable in columns 1-2 is a dummy for dropping out of middle school (Panel A) or high school (Panel B). The dependent variable in columns 3-4 is equal to one if the individual is currently working. The dependent variable in columns 5-8 is the log of household labor income per capita (col. 5-6), and per capita household labor income considering only income of members aged 25+ (cols 7-8). All columns restrict to individuals of middle school age (Panel A) or high school age (Panel B). All specifications include commuting zone, year-quarter, and quarter fixed effects, and control for sex, age, and household schooling of members aged 25+. Observations are weighted using normalized survey weights. Standard errors clustered at the commuting zone level are reported in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

⟷ Return to text

Table H.3: Robustness to minimum wage increases

	Dependent variable: School dropout			
	(1)	(2)	(3)	(4)
Panel A: Middle school age				
Exposure × Post	-0.012**	-0.006	-0.013*	-0.011+
	(0.005)	(0.010)	(0.005)	(0.006)
HH Income pc (log)				-0.000
				(0.002)
Num.Obs.	665100	42167	622933	494326
Sample	Full	ZLFN	Not ZLFN	Full
Panel B: High school age				
Exposure × Post	-0.023***	-0.016	-0.023***	-0.020*
	(0.007)	(0.015)	(0.007)	(0.010)
HH Income pc (log)				0.007+
				(0.004)
Num.Obs.	674437	43257	631180	494181
Sample	Full	ZLFN	Not ZLFN	Full

Notes: This table reports difference-in-differences estimates of the effect of exposure to clandestine drug labs on school dropout and household labor income. The dependent variable in columns 1-4 is a dummy for dropping out of high school (Panel A) or middle school (Panel B). The dependent variable in columns 5-7 is the log of household labor income per capita (col. 5), per capita household labor income excluding the individual's own income (col. 6), and per capita household labor income considering only income of members aged 25+ (col. 7). Column 1 uses the full sample. Column 2 restricts to municipalities in the Zona Libre de la Frontera Norte (ZLFN). Columns 3-4 exclude ZLFN municipalities, with column 4 additionally controlling for household income per capita. All columns restrict to individuals of high school age (Panel A) or middle school age (Panel B). All specifications include commuting zone, year-quarter, and quarter fixed effects, and control for sex, age, and household schooling of members aged 25+. Observations are weighted using normalized survey weights. Standard errors clustered at the commuting zone level are reported in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

⟷ Return to text

H.4 COVID-19

The pandemic began eleven months after the ban, but it does not explain the results. Schools closed nationwide in March 2020 by decision of the Ministry of Public Education, and students were automatically promoted and re-enrolled until in-person classes resumed in September 2021. The closures therefore did not differ between exposed and control areas.

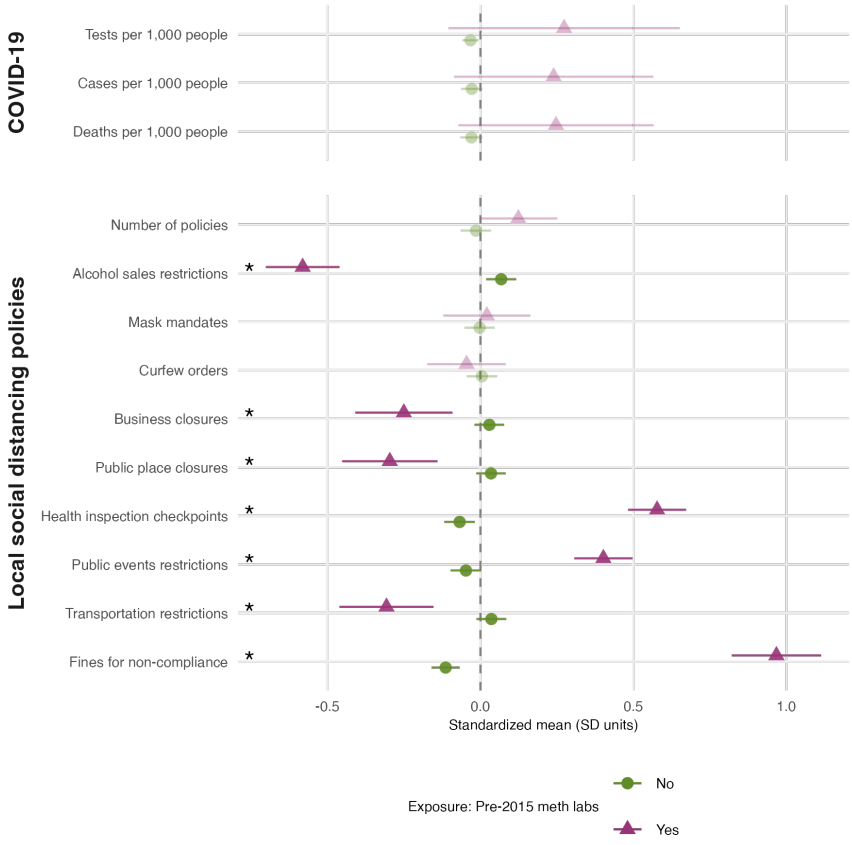
Exposed and control areas were similarly affected by the pandemic. Figure H.5 compares tests, cases, and deaths per capita in 2020 and local social distancing policies.[29] Cases and deaths do not differ. Exposed areas were less likely to adopt some local restrictions, such as alcohol sales bans and business closures, but the number of policies adopted is similar, and local governments did not decide school closures.

The effect of exposure holds in areas less affected by the pandemic. Table H.4 interacts exposure

[29]COVID-19 severity comes from administrative data of the Health Ministry (de Salud, 2025), and data on municipal social distancing policies from Enríquez et al. (2025).

with COVID-19 cases and deaths per capita (columns 1–2) and excludes the commuting zones with the most cases or deaths (columns 3–6). The interactions are small: at high school age, they are positive and significant but offset only a small part of the effect at the mean number of cases. When the hardest-hit commuting zones are excluded, the effect at high school age is 2.7 to 3.2 percentage points, larger than in the baseline. At middle school age the estimates are similar to the baseline but less precise.

Figure H.5: Balance on COVID-19 severity and local policy response by exposure to synthetic drug production



Notes: The figure assesses balance in COVID-19 severity and the local policy response across commuting zones by treatment status. Each point is the mean of the indicated measure (e.g., 2020 COVID-19 cases and deaths per capita and local policy indicators), standardized to mean zero and unit standard deviation, shown separately for treated commuting zones (any methamphetamine laboratory seized before 2015) and control zones; horizontal bars are 95% confidence intervals. Asterisks denote treated–control differences significant in a two-sample t-test.
Source: Author's calculations.

↩ Return to text

Table H.4: Heterogeneity by COVID-19 severity

	Interaction with		Excluding CZ with top			
	Cases	Deaths	Cases		Deaths	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Middle school dropout						
Exposure × Post	-0.010*	-0.009*	-0.017+	-0.012+	-0.008	-0.009
	(0.004)	(0.004)	(0.009)	(0.006)	(0.007)	(0.008)
Exposure × Post × Cases p.c.	0.001					
	(0.000)					
Exposure × Post × Deaths p.c.		0.003				
		(0.002)				
N	665,100	665,100	171,289	258,508	278,045	164,793
Panel B: High school dropout						
Exposure × Post	-0.022***	-0.021**	-0.030+	-0.030**	-0.027*	-0.032+
	(0.007)	(0.007)	(0.017)	(0.011)	(0.013)	(0.018)
Exposure × Post × Cases p.c.	0.001*					
	(0.000)					
Exposure × Post × Deaths p.c.		0.007*				
		(0.003)				
N	674,437	674,437	168,072	255,358	276,206	160,038

Notes: This table reports difference-in-differences estimates of the effect of increased synthetic drug production on school dropout. The dependent variable is an indicator for not being enrolled in school among individuals of the relevant school age: ages 12-14 for middle school (Panel A) and ages 15-17 for high school (Panel B). Exposure is an indicator for any methamphetamine laboratory seized in the commuting zone before 2015. Post is an indicator for periods after April 2019. Columns (1)-(2) interact the treatment effect with COVID-19 cases and deaths per 1,000 population in 2020, respectively. Columns (3)-(4) exclude commuting zones in the top 5% and 10% of COVID-19 cases. Columns (5)-(6) exclude commuting zones in the top 5% and 10% of COVID-19 deaths. All specifications include commuting zone and year fixed effects and control for gender, age, and household education. Standard errors clustered at the commuting zone level are reported in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

⟷ Return to text

H.5 Prospera rollback

The government ended the conditional cash transfer program Prospera in early 2019, two months before China's fentanyl ban. Prospera transferred money to poor households conditional on their children's school attendance and reached about 7 million households (Parker and Todd, 2017). Its replacement, the Becas Benito Juárez, reached households with delays and was less targeted to the poorest localities (Marquez-Padilla et al., 2025). Marquez-Padilla et al. (2025) show that the rollback reduced school enrollment, especially among boys of high school age. Prospera was more common in poor, rural areas, which are more often control areas in my sample, so the rollback could raise dropout in control areas and make dropout in exposed areas look relatively lower.

The effect of exposure remains when I allow the rollback to affect high-Prospera areas differently. Table H.5 interacts the post-ban period with indicators for commuting zones in the top 25%

or top 50% of Prospera coverage in 2017. Consistent with Marquez-Padilla et al. (2025), dropout rose after 2019 in high-coverage areas, by 5.8 percentage points at middle school age and 8.4 at high school age in the top 25%. In low-coverage areas, exposure still reduces dropout at high school age, by 1.7 and 1.2 percentage points depending on the definition. At middle school age the estimates are smaller (0.8 and 0.3 percentage points). Within the top 25% of coverage, the effect of exposure is larger than in low-coverage areas, so the rollback does not generate the result. Excluding the most rural commuting zones (columns 3–4), exposure reduces dropout by 0.9 to 1.1 percentage points at middle school age and 1.7 to 2.2 at high school age.

The timing of the effects also differs. The rollback reduced enrollment immediately, at the start of the 2019–2020 school year (Marquez-Padilla et al., 2025), while the effect of exposure grows over several years after the ban (Figure 3).

Table H.5: Robustness to Prospera rollback

	Dependent variable: School dropout			
	Interaction with		Excluding CZ with top	
	(1)	(2)	(3)	(4)
Panel A: Middle school age				
Exposure × Post	-0.008+	-0.003	-0.011*	-0.009+
	(0.004)	(0.004)	(0.005)	(0.005)
Post × Prospera Top 25%	0.058***			
	(0.015)			
Post × Prospera Top 50%		0.048***		
		(0.009)		
Exposure × Post × Prospera Top 25%	-0.111***			
	(0.036)			
Exposure × Post × Prospera Top 50%		-0.030		
		(0.024)		
Num.Obs.	665100	665100	654639	632029
Panel B: High school age				
Exposure × Post	-0.017**	-0.012+	-0.022***	-0.017*
	(0.006)	(0.006)	(0.007)	(0.007)
Post × Prospera Top 25%	0.084***			
	(0.022)			
Post × Prospera Top 50%		0.064***		
		(0.013)		
Exposure × Post × Prospera Top 25%	-0.186+			
	(0.100)			
Exposure × Post × Prospera Top 50%		-0.059		
		(0.042)		
Num.Obs.	674437	674437	665071	643988

Notes: This table reports difference-in-differences estimates of the effect of increased synthetic drug production on school dropout, testing robustness to the rollback of the Prospera conditional cash transfer program. The dependent variable is an indicator for not being enrolled in school among individuals of the relevant school age: ages 12–14 for middle school (Panel A) and ages 15–17 for high school (Panel B). Exposure is an indicator for any methamphetamine laboratory seized in the commuting zone before 2015. Post is an indicator for periods after April 2019. Columns (1)–(2) interact the treatment effect with indicators for high Prospera penetration, defined as commuting zones in the top 25th percentile (column 1) or top 50th percentile (column 2) of the 2017 beneficiary share distribution (ratio of enrolled households to total households from the 2010 census). Columns (3)–(4) exclude the most rural commuting zones, where Prospera penetration was highest. All specifications include commuting zone and year fixed effects and control for gender, age, and household education. Standard errors clustered at the commuting zone level are reported in parentheses + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

↩ Return to text

H.6 Assignment of tertiary enrollment to origin commuting zones

This section shows that the tertiary enrollment results do not depend on the migration matrix used to assign enrollment to origins. I first re-estimate the main specification with three alternatives to the baseline matrix. I then discuss what the origin estimates add compared to estimates by campus

location, report the attenuation factor described in Appendix D.2, and compare the results with university school attendance using the survey ENOE, which does not use the matrix.

Alternative matrix assignments. I re-estimate the main specification with three alternative ways of assigning enrollment (Table H.6). The first uses separate migration matrices for women and men. The baseline uses one matrix for both, which assumes that women and men from the same origin study in the same destinations in the same proportions. With separate matrices, the enrollment of each sex is assigned with its own migration pattern. The second uses a matrix built in the same way from the 2020 Census, with origin given by the commuting zone of residence in 2015. This matrix is measured after the ban, so it partly reflects where students studied after the shock (Figure D.1). The third does not assign enrollment to origins. It measures enrollment in the commuting zone where each campus is located and uses that commuting zone's exposure. This estimate does not use any matrix, so it relies on neither of the two assumptions in Appendix D.2.

All four sets of estimates are nearly identical. For chemistry programs, the campus-level estimate is 0.058, compared with 0.057 in the baseline; for women it is 0.062 compared with 0.061, and for men 0.049 compared with 0.048. The separate matrices by sex and the 2020 matrix change no coefficient by more than 0.002. Figure H.6 shows that the event studies for chemistry programs also coincide year by year. The increase in chemistry enrollment is therefore present in the campus enrollment data before any assignment to origins.

Contribution of origin-level estimates. The origin estimates measure the response of young people from exposed areas, which the campus-level estimate cannot separate from exposed areas attracting students from elsewhere. An increase in enrollment on campuses in treated areas can come from young people from those areas enrolling more or choosing different fields. It can also come from treated areas attracting students from control areas, for example because new programs open there or because the labor market opportunities there are more attractive. The campus-level estimate combines the two. The origin and campus-level estimates coincide because few students study outside their home commuting zone: about 95% of the students on campuses in treated commuting zones come from treated origins. Reading the result as a change in the enrollment of young people from treated origins, and not as treated areas attracting students, relies on Assumption 1. The 2010 and 2020 Censuses support this assumption. The share of students from control origins who study in a treated commuting zone fell from 3.1% to 2.7% between the two censuses, and the share of students from treated origins who study in a control commuting zone fell from 1.4% to 1.1% (Figure D.1).

Attenuation factor. The imputation makes the estimates about 10% smaller than the true effects. When Assumption 1 holds, it shrinks them by the factor $\lambda = \bar{T}_1 - \bar{T}_0$ derived in Appendix D.2. The factor depends only on the 2010 matrix and on enrollment before 2019, so I compute it directly from the data. On average, 94% of the classmates of students from treated origins also come from treated origins, compared with 5% for students from control origins. The attenuation factor is therefore about 0.90 in every field. This matters little for the conclusions in this case: the imputation does not change the sign of an effect, and the evidence below shows no increase in total enrollment.

Standard errors. Standard errors clustered by state are larger, and with them the estimate for men is no longer significant. Origins that send students to the same campuses have related outcomes, which standard errors clustered by commuting zone do not account for. Table H.6 also reports standard errors clustered by state, in brackets. They are larger for chemistry programs (0.030 instead of 0.023), and the p-values become 0.058 for all students, 0.046 for women and 0.115 for men.

Did more young people enroll in university? The evidence that the shock raised university enrollment in treated areas is limited. Total enrollment does not change in Table 3, but its event study shows a relative decline in treated areas before the ban (Figure 5), so it does not settle the question on its own. ENOE allows a second measure that does not use the matrix. It records whether young people attend school in the commuting zone where they live. I use individuals of tertiary age, the four years after the expected start of university, and define the outcome as attending school after completing at least 12 years of schooling, which approximates attending university. The specification is the same as for school dropout in Section 6: controls for sex, age and household schooling, fixed effects for commuting zone, quarter-year and quarter, survey weights, and standard errors clustered by commuting zone.

School attendance at tertiary age does not change in exposed areas. Table H.7 reports the estimates: the estimate is -0.003 (s.e. 0.007), against a control mean of 0.25, and it is close to zero for both women and men. The estimates are the same when the sample is limited to commuting zones with a campus, which contain 89% of young people of tertiary age. Figure H.7 shows the event study. The coefficients are close to zero before the ban. After the ban, attendance falls during the school closures and returns to its earlier level by the third year. ENOE records young people where they live, so students who moved to study are observed in their destination, and it does not record the field of study of current students.

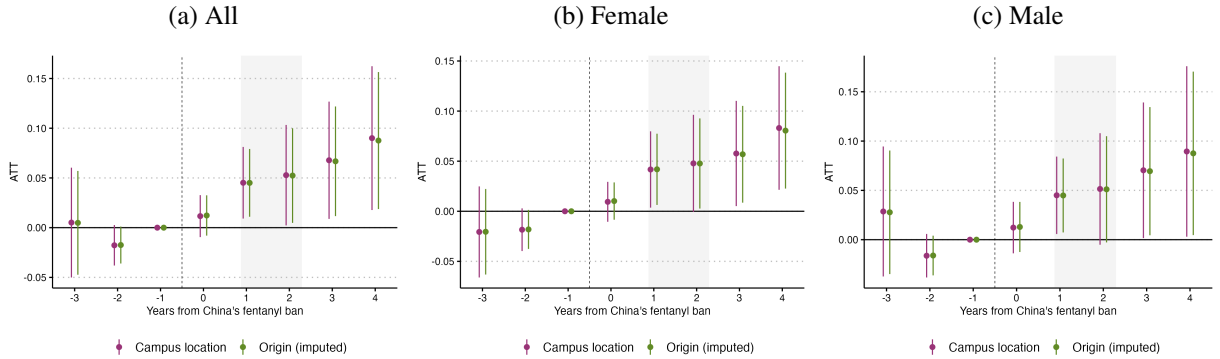
Together, these results indicate that the shock did not raise the number of young people who enroll in university in treated areas. The increase in chemistry enrollment therefore reflects mainly a change in what young people study.

Table H.6: Tertiary enrollment: alternative assignments of enrollment to commuting zones

	All fields	Chemistry	Non-chemistry	Non-STEM+H
	(1)	(2)	(3)	(4)
Panel A. All				
Origin, 2010 matrix (baseline)	0.000	0.057*	-0.016	-0.028
	(0.022)	(0.023)	(0.020)	(0.026)
	[0.023]	[0.030]	[0.022]	[0.026]
Origin, 2010 matrices by sex	0.001	0.058*	-0.016	-0.027
	(0.022)	(0.023)	(0.020)	(0.026)
Origin, 2020 matrix	-0.001	0.057*	-0.018	-0.030
	(0.022)	(0.023)	(0.020)	(0.026)
Campus location, no imputation	-0.002	0.058*	-0.018	-0.031
	(0.023)	(0.025)	(0.022)	(0.027)
Panel B. Female				
Origin, 2010 matrix (baseline)	-0.009	0.061**	-0.049+	-0.029
	(0.020)	(0.022)	(0.027)	(0.026)
	[0.023]	[0.030]	[0.029]	[0.027]
Origin, 2010 matrices by sex	-0.010	0.061**	-0.049+	-0.030
	(0.021)	(0.022)	(0.027)	(0.026)
Origin, 2020 matrix	-0.011	0.061**	-0.050+	-0.031
	(0.021)	(0.022)	(0.027)	(0.026)
Campus location, no imputation	-0.012	0.062*	-0.051+	-0.033
	(0.022)	(0.024)	(0.030)	(0.027)
Panel C. Male				
Origin, 2010 matrix (baseline)	0.010	0.048+	0.010	-0.024
	(0.024)	(0.025)	(0.019)	(0.027)
	[0.025]	[0.031]	[0.019]	[0.026]
Origin, 2010 matrices by sex	0.011	0.049+	0.011	-0.023
	(0.024)	(0.026)	(0.019)	(0.027)
Origin, 2020 matrix	0.009	0.048+	0.008	-0.026
	(0.024)	(0.025)	(0.019)	(0.027)
Campus location, no imputation	0.008	0.049+	0.008	-0.027
	(0.025)	(0.027)	(0.020)	(0.028)

Notes: Poisson estimates of Exposure × Post, with commuting zone and year fixed effects, for ANUIES enrollment in 2016–2023. Columns follow Table 3. "Origin, 2010 matrix" is the baseline specification. "Origin, 2010 matrices by sex" assigns women's and men's enrollment with separate matrices built from students of each sex; total enrollment is the sum of the two. "Origin, 2020 matrix" builds the matrix from the 2020 Census, with origin given by residence in 2015. "Campus location" measures enrollment in the commuting zone where each campus is located and uses that commuting zone's exposure. Standard errors clustered by commuting zone in parentheses; for the baseline, standard errors clustered by state in brackets. Significance stars refer to standard errors clustered by commuting zone. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.
Source: ANUIES and Mexican Population Censuses 2010 and 2020; author's calculations.

Figure H.6: Chemistry enrollment: origin and campus-level event studies



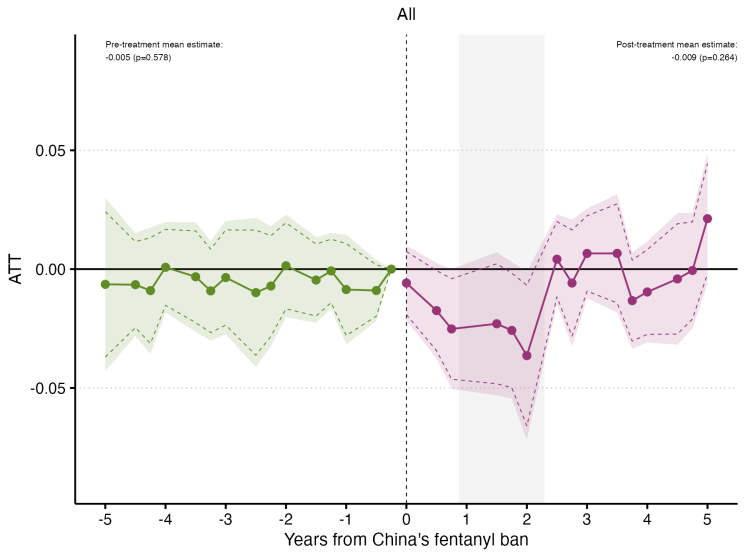
Notes: Poisson event-study estimates of the effect of exposure on enrollment in STEM+H programs with chemistry content, relative to 2018, with commuting zone and year fixed effects. "Origin (imputed)" is the baseline, with enrollment assigned to origin commuting zones with the 2010 migration matrix. "Campus location" measures enrollment in the commuting zone where each campus is located. Bars are 95% confidence intervals from standard errors clustered by commuting zone. The shaded area marks the school closures.
Source: ANUIES and Mexican Population Census 2010; author's calculations.

Table H.7: School attendance at tertiary age (ENOE)

	All commuting zones			Commuting zones with a campus		
	All	Female	Male	All	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure × Post	-0.003	-0.009	0.003	-0.003	-0.008	0.003
	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)	(0.009)
Control mean	0.252	0.265	0.238	0.279	0.295	0.264
Commuting zones	1,203	1,184	1,187	486	485	485
Observations	750,520	373,378	377,142	704,353	350,205	354,148

Notes: Linear probability estimates of Exposure × Post. The outcome equals one for individuals who attend school and have completed at least 12 years of schooling. The sample consists of individuals of tertiary age (the four years after the expected start of university) in ENOE, 2013–2024, excluding third quarters. Columns (4)–(6) restrict the sample to commuting zones with ANUIES enrollment before 2019. All specifications control for sex, age and household schooling, include commuting zone, quarter-year and quarter fixed effects, and use survey weights. Post equals one from the second quarter of 2019. Standard errors clustered by commuting zone in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01.
Source: ENOE and ANUIES; author's calculations.

Figure H.7: School attendance at tertiary age (ENOE): event study



Notes: Event-study estimates of the effect of exposure on the probability of attending school with at least 12 years of schooling completed, for individuals of tertiary age. Quarterly event time relative to the first quarter of 2019; third quarters are excluded. Linear probability model with controls for sex, age and household schooling, fixed effects for commuting zone, quarter-year and quarter, and survey weights. The shaded band is the 95% confidence interval and the dashed lines the 90% interval, from standard errors clustered by commuting zone. The grey area marks the school closures.
Source: ENOE; author's calculations.

I Validation of Exposure to Synthetic Drug Production Measures Using Byproducts in Water

I.1 Background: Chemical Signatures of Methamphetamine Production

Clandestine drug laboratories generate substantial environmental contamination through the discharge of toxic waste into surrounding ecosystems. A 2024 investigation by Quinto Elemento Lab documented that laboratories in Mexico routinely extract water from rivers and irrigation canals for drug synthesis, then discharge contaminated wastewater—containing acids, solvents, and heavy metals—back into the same water bodies (Lab, 2025). Environmental activists in Sinaloa report finding laboratory equipment scattered in rivers, carried downstream by rainfall from mountain production sites. The investigation found that Mexican authorities lack protocols for environmental remediation: when military forces dismantle laboratories in remote areas, they typically destroy chemicals on-site through open-air burning or dumping rather than proper disposal, extending contamination beyond the original site.

Methamphetamine production generates distinctive environmental contamination. The dominant synthesis method in Mexico since approximately 2012 is the P2P (phenyl-2-propanone) reduction method, which replaced the earlier Birch reduction method after Mexican authorities restricted pseudoephedrine imports beginning in 2008.

The P2P method involves reducing phenyl-2-propanone using an aluminum-mercury amalgam catalyst, then reacting the intermediate with methylamine. This process generates several chemical byproducts:

- **Mercury:** The aluminum-mercury amalgam catalyst releases mercury into wastewater. Unlike most industrial sources of mercury (gold mining, coal combustion), the concentrations and geographic distribution of mercury from drug production should correlate specifically with laboratory locations.
- **Toluene:** Used as a solvent in multiple stages of synthesis and purification. Toluene is also used industrially (in paints, adhesives, and gasoline), but its co-occurrence with mercury in areas lacking other industrial activity provides a distinctive signature.
- **Acidic/basic waste:** Production involves alternating acidification and basification steps, generating corrosive wastewater that may affect pH and dissolved oxygen levels.

I.2 Data Construction

I.2.1 Water Quality Monitoring Data

I use data from Mexico's National Water Quality Monitoring Network (RENAMECA), administered by the National Water Commission (CONAGUA). RENAMECA conducts systematic water quality monitoring at approximately 5,000 sites across Mexico's major surface water bodies, including rivers, lakes, reservoirs, and coastal zones. Monitoring frequency varies by site, with most sites sampled quarterly or annually.

For each site-year, I construct the following pollution indices:

1. **Mercury detection rate:** Share of samples with mercury (Hg) concentrations above the detection limit.

2. **Toluene detection rate:** Share of samples with toluene concentrations above the detection limit.

3. **COD/BOD ratio:** Ratio of chemical oxygen demand to biological oxygen demand, standardized across the sample. Higher values indicate non-biodegradable industrial contamination; whereas values below 2 indicate organic/biological contamination (sewage, agricultural runoff).

I.2.2 Aggregation to Commuting Zones

I aggregate site-level measurements to commuting zones using inverse-distance weights. For each monitoring site, I calculate the distance to the centroid of all commuting zones within 100 kilometers. The weight assigned to site s for commuting zone c is:

$$w_{sc} = \frac{1/d_{sc}^2}{\sum_{s'} 1/d_{s'c}^2} \tag{49}$$

where d_{sc} is the distance from site s to the centroid of commuting zone c, and the sum is over all sites within 100 km of the commuting zone.

For commuting zone c in year t, the pollution index is:

$$\text{Pollution}_{ct} = \sum_{s} w_{sc} \cdot \text{Pollution}_{st} \tag{50}$$

This approach assigns greater weight to nearby monitoring sites while incorporating information from the broader watershed.

I.3 Empirical Specification

I estimate the relationship between laboratory exposure and water quality using:

$$\text{Pollution}_{ct} = \beta \cdot \text{Exposure}_c + \gamma \cdot X_{ct} + \mu_s + \tau_t + \varepsilon_{ct} \tag{51}$$

where:

- Pollution_{ct} is a pollution index for commuting zone c in year t
- Exposure_c is the laboratory exposure measure (either observed pre-2015 seizures or machine learning predictions)
- X_{ct} includes the COD/BOD ratio (to control for general industrial activity) and the number of monitoring sites (to account for measurement precision)
- μ_s are state fixed effects
- τ_t are year fixed effects

I weight observations by the sum of inverse-distance weights (reflecting monitoring coverage) and report standard errors clustered at the commuting zone level.

I.4 Results

The estimates from this validation exercise are reported in Table 1 in Section 7.

J Violence and survey non-response

This appendix reports the estimates underlying the measurement premises of Section 8: the homicide results (§8.1) and the survey non-response results (§8.3).

J.1 Homicides

Table J.1 reports the homicide difference-in-differences across specifications, and Figure J.1 shows the corresponding event studies. The point estimates are uniformly small and, if anything, positive; none supports a differential decline in violence, and the per-capita rate result does not depend on whether population is measured with annual projections or fixed to the five-year census anchors.

Table J.1: Exposure to synthetic drug production and intentional homicides

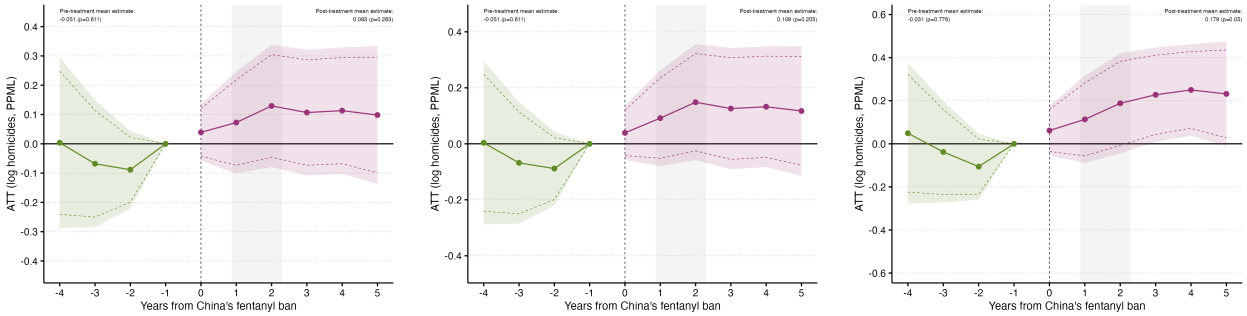
	Rate (annual)	Rate (census)	Count (unweighted)	High-violence CZs
Exposure × Post	0.134	0.131	0.148	0.209+
	(0.100)	(0.101)	(0.107)	(0.114)
N	17,765	17,765	17,765	7,590
Mean of dep. var.	17.793	17.793	9.686	26.716

Notes: Poisson pseudo-maximum-likelihood estimates at the commuting-zone × year level, 2015–2024, with commuting-zone and year fixed effects; standard errors clustered by commuting zone. The dependent variable is the intentional-homicide count. "Rate (annual)" and "Rate (census)" add a log-population offset so the coefficient is a proportional change in the per-capita rate, using annual CONAPO population and population fixed to the five-year census/intercensal anchors (2015, 2020), respectively. "Count (unweighted)" applies no population adjustment. "High-violence CZs" restricts to commuting zones above the median pre-period (2015–2018) homicide rate among those with any pre-period homicide. "Mean of dep. var." is the control homicide rate per 100,000 for the rate and high-violence columns and the control mean count for the unweighted column. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

↩ Return to text

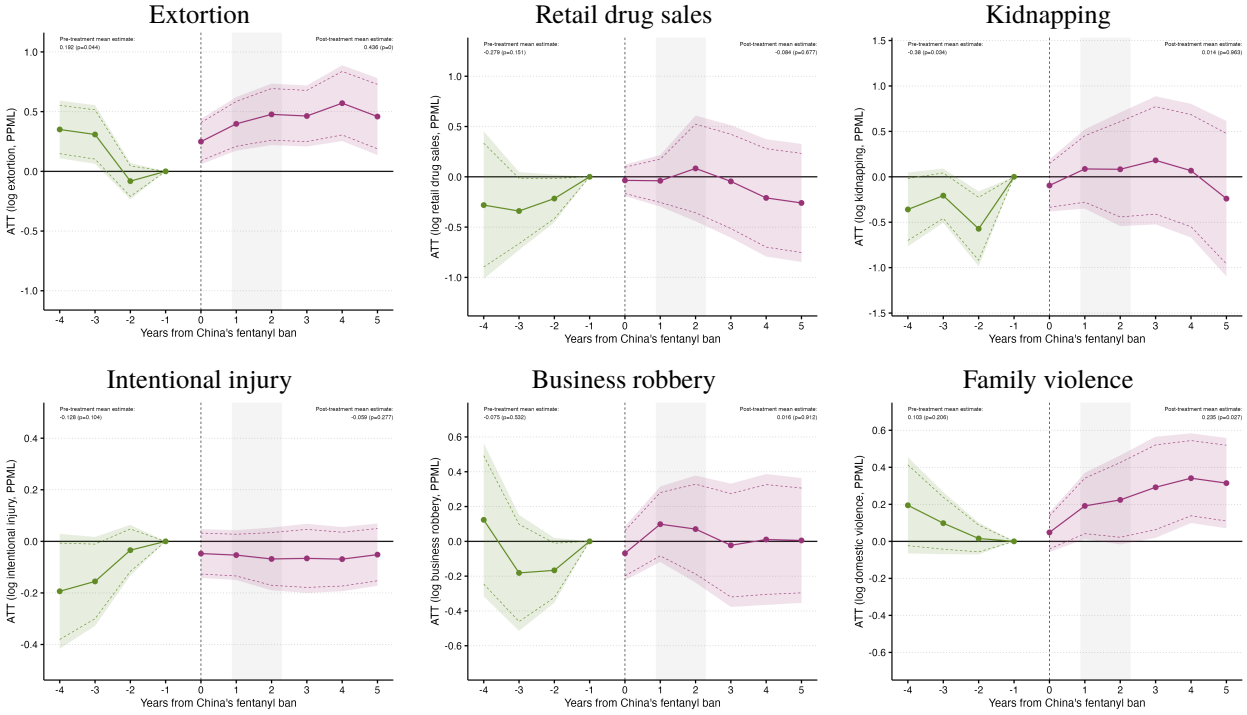
Figure J.1: Exposure and intentional homicides: robustness event studies

(a) Rate, census-anchored population (b) Unweighted count (c) Rate, high-violence CZs



Notes: Event studies corresponding to the specifications in Table J.1, relative to the year before the ban ($t = -1$). Poisson pseudo-maximum-likelihood with commuting-zone and year fixed effects; standard errors clustered by commuting zone. Panel (a) uses a log-population offset with population fixed to the five-year census anchors; panel (b) is the unweighted homicide count; panel (c) is the per-capita rate estimated on commuting zones above the median pre-period homicide rate.

Figure J.2: Exposure and other crimes



Notes: Event studies of crime rates at the commuting-zone × year level, 2015–2025, with the same design as Figure 7: Poisson pseudo-maximum-likelihood with a log-population offset, commuting-zone and year fixed effects, and reference year $t = -1$. Bands are 95% confidence intervals from standard errors clustered at the commuting zone level.
Source: Author's calculations based on SESNSP crime incidence data (new methodology) and CONAPO population projections.

↩ Return to text

J.2 Survey non-response

Table J.2 reports difference-in-differences estimates of exposure on survey non-response. Panel A is unit non-response at the dwelling × year level: the share of inhabited dwellings that are outright refusals, refusals plus "other reason," and any non-complete visit. Panel B is item non-response among employed respondents: missing income, income heaping at round values, and missing occupation. Exposure raises outright refusals but leaves item non-response essentially unchanged, indicating that whole households or workers opt out rather than respondents concealing a single attribute.

Table J.2: Exposure to synthetic drug production and survey non-response

Panel A. Unit non-response (dwelling × year)

	Refusals only	Refusals + Otro motivo	Any non-complete
Exposure × Post	0.007*	0.012	0.018
	(0.003)	(0.012)	(0.013)
N	5,489,535	5,489,535	5,489,535
Control mean	0.027	0.052	0.078

Panel B. Item non-response among employed respondents

	Income missing	Heap at 1,000	Heap at 5,000	Occupation missing
Exposure × Post	0.015	-0.009+	-0.003	-0.001
	(0.014)	(0.005)	(0.003)	(0.001)
N	7,554,458	5,574,481	5,574,481	7,893,739
Control mean	0.258	0.205	0.054	0.001

Notes: Difference-in-differences estimates; standard errors clustered by commuting zone in parentheses. Both panels are at the year level with commuting-zone, year, and quarter fixed effects and commuting-zone linear trends. *Exposure* equals one for commuting zones with a methamphetamine laboratory seized before 2015; *Post* equals one from 2019 (China's fentanyl ban). "Control mean" is the mean of the dependent variable in the untreated (*Exposure* = 0 or pre-ban) sample.
Panel A is estimated on all inhabited dwellings and is unweighted, since ENOE expansion factors are zero for non-respondents by construction. Each outcome is a 0/1 indicator for the dwelling's disposition: *Refusals only* equals one if the occupants refused the interview; *Refusals + Otro motivo* adds dwellings not interviewed for "another reason (dwelling inhabited)"; *Any non-complete* equals one for any inhabited dwelling that did not yield a completed interview.
Panel B is estimated on employed respondents, weighted by survey weights, and controls for sex, age, and household adult schooling. *Income missing* is the share of paid workers reporting zero or missing labor income. *Heap at 1,000* and *Heap at 5,000* are the share of *positive* reported incomes that are exact multiples of 1,000 and 5,000 pesos—a proxy for rounded, imprecise income reporting (higher values indicate less precise reports). *Occupation missing* is the share of workers with an unspecified occupation (SINCO code 9999). + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

↩ Return to text

References

Adukia, Anjali, Sam Asher, and Paul Novosad, "Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction," *American Economic Journal: Applied Economics*, January 2020, *12* (1), 348–376.

Alesina, Alberto, Salvatore Piccolo, and Paolo Pinotti, "Organized Crime, Violence, and Politics," *The Review of Economic Studies*, March 2019, *86* (2), 457–499.

Alpert, Abby, David Powell, and Rosalie Liccardo Pacula, "Supply-Side Drug Policy in the Presence of Substitutes: Evidence from the Introduction of Abuse-Deterrent Opioids," *American Economic Journal: Economic Policy*, 2018, *10* (4), 1–35.

Asmann, Parker, "Methamphetamine Taking Over Mexico's Domestic Drug Market," April 2021.

Athey, Susan and Guido W Imbens, "The state of applied econometrics: Causality and policy evaluation," *Journal of Economic Perspectives*, 2017, *31* (2), 3–32.

_ and _, "Machine learning methods that economists should know about," *Annual Review of Economics*, 2019, *11*, 685–725.

Atkin, David, "Endogenous Skill Acquisition and Export Manufacturing in Mexico," *American Economic Review*, August 2016, *106* (8), 2046–2085.

Battagliola, Maria Laura and others, "Refining public policies with machine learning: The case of tax auditing," *Journal of Econometrics*, 2024.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76* (2), 169–217. Publisher: University of Chicago Press.

Bell, Brian, Rui Costa, and Stephen Machin, "Why Does Education Reduce Crime?," *Journal of Political Economy*, March 2022, *130* (3), 732–765. Publisher: The University of Chicago Press.

Blattman, Christopher, Gustavo Duncan, Benjamin Lessing, and Santiago Tobón, "Gang Rule: Understanding and Countering Criminal Governance," *The Review of Economic Studies*, 2025, *92* (3), 1497–1531.

Blom, Erica, Brian C. Cadena, and Benjamin J. Keys, "Investment over the Business Cycle: Insights from College Major Choice," *Journal of Labor Economics*, 2021, *39* (4), 1043–1082.

Bound, John, Breno Braga, Joseph M. Golden, and Gaurav Khanna, "Recruitment of Foreigners in the Market for Computer Scientists in the United States," *Journal of Labor Economics*, 2015, *33* (S1), S187–S223.

Brown, Ryan and Andrea Velásquez, "The effect of violent crime on the human capital accumulation of young adults," *Journal of Development Economics*, 2017, *127*, 1–12.

Calderón, Mariana, Josué Cortés, Jorge Pérez Pérez, and Alejandrina Salcedo, "Disentangling the Effects of Large Minimum Wage and VAT Changes on Prices: Evidence from Mexico," *Labour Economics*, January 2023, *80*, 102294.

Campaniello, Nadia, Rowena Gray, and Giovanni Mastrobuoni, "Returns to Education in Criminal Organizations: Did Going to College Help Michael Corleone?," *Economics of Education Review*, 2016, *54*, 242–258.

Campedelli, Gian Maria, Gianmarco Daniele, and Marco Le Moglie, "Mafia, Politics and Machine Predictions," CEPR Discussion Paper DP19322, Centre for Economic Policy Research 2024.

Castillo, Juan Camilo and Dorothy Kronick, "The Logic of Violence in Drug War," *American Political Science Review*, August 2020, *114* (3), 874–887.

_ , Daniel Mejía, and Pascual Restrepo, "Scarcity without Leviathan: The Violent Effects of Cocaine Supply Shortages in the Mexican Drug War," *Review of Economics and Statistics*, 2020, *102* (2), 269–286.

Chang, Eunsik and Maria Padilla-Romo, "When crime comes to the neighborhood: Short-term shocks to student cognition and secondary consequences," *Journal of Labor Economics*, 2023, *41* (4), 997–1039.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo, "Housing Booms and Busts, Labor Market Opportunities, and College Attendance," *American Economic Review*, October 2018, *108* (10), 2947–2994.

Congressional Research Service, "Illicit Fentanyl and China's Role," CRS In Focus IF10890, Congressional Research Service 2019.

Daniele, Gianmarco, Marco Le Moglie, and Federico Masera, "Pains, guns and moves: The effect of the U.S. opioid epidemic on Mexican migration," *Journal of Development Economics*, January 2023, *160*, 102983.

de Roux, Daniel, Boris Perez, Andrés Moreno, Maria del Pilar Villamil, and César Figueroa, "Tax Fraud Detection for Under-Reporting Declarations Using an Unsupervised Machine Learning Approach," in "Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining" KDD '18 Association for Computing Machinery New York, NY, USA July 2018, pp. 215–222.

de Salud, Secretaría, "Dirección General de Epidemiología - Datos Abiertos Bases Históricas," 2025.

Delavande, Adeline and Basit Zafar, "University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints," *Journal of Political Economy*, October 2019, *127* (5), 2343–2393. Publisher: The University of Chicago Press.

Dell, Melissa, "Trafficking Networks and the Mexican Drug War," *American Economic Review*, June 2015, *105* (6), 1738–1779.

Dix-Carneiro, Rafael, Rodrigo R. Soares, and Gabriel Ulyssea, "Economic Shocks and Crime: Evidence from the Brazilian Trade Liberalization," *American Economic Journal: Applied Economics*, 2018, *10* (4), 158–195.

Dube, Oeindrila and Juan F. Vargas, "Commodity Price Shocks and Civil Conflict: Evidence from Colombia," *The Review of Economic Studies*, 2013, *80* (4), 1384–1421.

Dudley, Steven, "El flujo de precursores químicos para la producción de drogas sintéticas en México," Technical Report, InSight Crime 2023.

Duranton, Gilles, "A Proposal to Delineate Metropolitan Areas in Colombia," *Revista Desarrollo y Sociedad*, January 2015, (75), 223–264. Number: 75.

Ehrlich, Isaac, "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy*, 1973, *81* (3), 521–565. Publisher: The University of Chicago Press.

Enríquez, José Ramón, Horacio Larreguy, and Alberto Simpser, "The impact of confounders, spillovers and interactions on social distancing policy effects estimates," *Scientific Reports*, October 2025, *15* (1), 34350. Publisher: Nature Publishing Group.

Europol, ed., *European Union Serious and Organised Crime Threat Assessment: The Changing DNA of Serious and Organised Crime*, Luxembourg: Publications Office of the European Union, 2025.

Evans, William N., Ethan M. J. Lieber, and Patrick Power, "How the Reformulation of OxyContin Ignited the Heroin Epidemic," *The Review of Economics and Statistics*, 2019, *101* (1), 1–15.

Federal Bureau of Investigation, "2025 IC3 Annual Report," Technical Report, Internet Crime Complaint Center 2025.

Freeman, Richard B., "A Cobweb Model of the Supply and Starting Salary of New Engineers," *Industrial and Labor Relations Review*, 1976, *29* (2), 236–248.

Friedman, Joseph and Chelsea L. Shover, "Charting the fourth wave: Geographic, temporal, race/ethnicity and demographic trends in polysubstance fentanyl overdose deaths in the United States, 2010–2021," *Addiction*, 2023, *118* (12), 2477–2485.

Garcia, Sara, "Químicos y cocineros: así mata el fentanilo," May 2023.

Grogger, Jeff, "Market wages and youth crime," *Journal of labor Economics*, 1998, *16* (4), 756–791.

Haanwinckel, Daniel, "Does regional variation in wage levels identify the effects of a national minimum wage?," April 2025. arXiv:2307.01284 [econ].

International Narcotics Control Board, "Precursors and Chemicals Frequently Used in the Illicit Manufacture of Narcotic Drugs and Psychotropic Substances: Report of the International Narcotics Control Board for 2022," Technical Report E/INCB/2022/4, United Nations, Vienna 2023.

INTERPOL, "INTERPOL Issues Global Warning on Human Trafficking-Fueled Fraud," 2023. Press release, 7 June 2023.

_ , "INTERPOL Releases New Information on Globalization of Scam Centres," 2025. Crime Trend Update, 30 June 2025.

Jarillo, Brenda, Beatriz Magaloni, Edgar Franco, and Gustavo Robles, "How the Mexican Drug War affects kids and schools? Evidence on effects and mechanisms," *International Journal of Educational Development*, 2016, *51*, 135–146.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer, "Prediction policy problems," *American Economic Review*, 2015, *105* (5), 491–495.

Lab, Quinto Elemento, "Narco Contaminación: La catástrofe ambiental de los laboratorios clandestinos," 2025.

Leo, Lorenzo Aldeco, Andrés Jurado, and Aurora A. Ramírez-Álvarez, "Internal migration and drug violence in Mexico," *Journal of Development Economics*, October 2024, *171*, 103334.

Lindo, Jason M and María Padilla-Romo, "Kingpin approaches to fighting crime and community violence: evidence from Mexico's drug war," *Journal of health economics*, 2018, *58*, 253–268.

Lochner, Lance and Enrico Moretti, "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *American Economic Review*, March 2004, *94* (1), 155–189.

Machin, Stephen and Matteo Sandi, "Crime and education," *Annual Review of Economics*, August 2025, *17* (1), 241–260. Num Pages: 20.

_ , Olivier Marie, and Sunčica Vujić, "The Crime Reducing Effect of Education," *The Economic Journal*, 2011, *121* (552), 463–484. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0297.2011.02430.x.

Marquez-Padilla, Fernanda, Susan Parker, and Tom Vogl, "Rolling Back Progresa: School and Work After the End of a Landmark Anti-Poverty Program," Technical Report w33527, National Bureau of Economic Research, Cambridge, MA February 2025.

Mejía, Daniel and Pascual Restrepo, "The Economics of the War on Illegal Drug Production and Trafficking," *Journal of Economic Behavior & Organization*, 2016, *126*, 255–275.

Melnikov, Nikita, Carlos Schmidt-Padilla, and Maria Micaela Sviatschi, "Gangs, labor mobility, and development," *Econometrica*, 2025, *93* (6), 2083–2121.

Mirenda, Litterio, Sauro Mocetti, and Lucia Rizzica, "The Economic Effects of Mafia: Firm Level Evidence," *American Economic Review*, 2022.

Monteiro, Joana and Rudi Rocha, "Drug battles and school achievement: evidence from Rio de Janeiro's favelas," *Review of Economics and Statistics*, 2017, *99* (2), 213–228.

Mullainathan, Sendhil and Jann Spiess, "Machine learning: An applied econometric approach," *Journal of Economic Perspectives*, 2017, *31* (2), 87–106.

National Drug Intelligence Center, "Gulf Coast High Intensity Drug Trafficking Area Drug Market Analysis 2009," Drug Market Analysis, U.S. Department of Justice, National Drug Intelligence Center 2009.

National Institute on Drug Abuse, "Fentanyl DrugFacts," 2021.

Parker, Susan W. and Petra E. Todd, "Conditional Cash Transfers: The Case of Progresa/Oportunidades," *Journal of Economic Literature*, September 2017, *55* (3), 866–915.

Pinotti, Paolo, "The Economic Costs of Organised Crime: Evidence from Southern Italy," *The Economic Journal*, 2015, *125* (586), F203–F232. Publisher: [Royal Economic Society, Wiley].

Prieto-Curiel, Rafael, Gian Maria Campedelli, and Alejandro Hope, "Reducing cartel recruitment is the only way to lower violence in Mexico," *Science*, 2023, *381* (6664), 1312–1316.

Quinones, Sam, *The Least of Us: True Tales of America and Hope in the Time of Fentanyl and Meth*, New York: Bloomsbury Publishing, 2021.

Ryoo, Jaewoo and Sherwin Rosen, "The Engineering Labor Market," *Journal of Political Economy*, 2004, *112* (S1), S110–S140.

Shah, Manisha and Bryce Millett Steinberg, "Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital," *Journal of Political Economy*, April 2017, *125* (2), 527–561. Publisher: The University of Chicago Press.

Silva, J. M. C. Santos and Silvana Tenreyro, "The Log of Gravity," *The Review of Economics and Statistics*, November 2006, *88* (4), 641–658.

Sobrino, Fernanda, "Mexican cartel wars: Fighting for the US opioid market," 2019.

Sviatschi, Maria Micaela, "Making a NARCO: Childhood Exposure to Illegal Labor Markets and Criminal Life Paths," *Econometrica*, 2022, *90* (4), 1835–1878. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17082.

United Nations Office on Drugs and Crime, "Three Precursors of the Most Common Synthesis Routes Used in Illicit Fentanyl Manufacture Now Under International Control," UNODC Laboratory and Scientific Service, Early Warning Advisory April 2022.

_ , "Crime-related illicit financial flows: latest progress," Technical Report, United Nations Office on Drugs and Crime (UNODC), Vienna 2023.

U.S. Department of the Treasury, "Treasury Sanctions Illicit Fentanyl Supply Network Supporting the Sinaloa Cartel," October 2025. Press Release No. SB0272ç.

U.S. Drug Enforcement Administration, "Fentanyl Flow in the United States," Intelligence Report DEA-DCT-DIR-008-20, U.S. Drug Enforcement Administration 2020.

_ , "DEA Laboratory Testing Reveals that 6 out of 10 Fentanyl-Laced Fake Prescription Pills Now Contain a Potentially Lethal Dose of Fentanyl," 2022.

U.S. Immigration and Customs Enforcement, "United States–Mexico Bi-National Criminal Proceeds Study," Cornerstone Report, U.S. Immigration and Customs Enforcement (ICE) 2010.

U.S. Senate Permanent Subcommittee on Investigations, "Combatting the Opioid Crisis: Exploiting Vulnerabilities in International Mail," Staff Report, Committee on Homeland Security and Governmental Affairs, United States Senate 2018.

Weinstein, Russell, "Local Labor Markets and Human Capital Investments," *Journal of Human Resources*, 2022, *57* (5), 1498–1525.

Zee, Art Van, "The Promotion and Marketing of OxyContin: Commercial Triumph, Public Health Tragedy," *American Journal of Public Health*, 2009, *99* (2), 221–227.