Title: Do We Pay for Innovation? The Rising Returns to Pharmacological Novelty in the U.S. Pharmaceutical Market

Authors:

Boshen Jiao, PhD[1,2]; Ian S. Haworth, PhD[3]; Darius N. Lakdawalla, PhD[1,2,4]

Affiliations:

1. Department of Pharmaceutical and Health Economics, Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California
2. Leonard D. Schaeffer Center for Health Policy & Economics, University of Southern California
3. Department of Pharmacology and Pharmaceutical Sciences, Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California
4. Department of Health Policy and Management, Sol Price School of Public Policy, University of Southern California

Corresponding author:

Darius N. Lakdawalla

Email: dlakdawa@usc.edu

ABSTRACT

Objective

To compare trends in US prescription drug revenue growth by drugs' pharmacological novelty, measured across three dimensions: molecular structure, therapeutic target, and delivery properties.

Design

Retrospective observational study linking pharmacological novelty to drug-level revenue, utilization, and rebates.

Setting

US outpatient prescription drug market.

Participants

FDA-approved branded new molecular entities marketed in the U.S. between 2000 and 2019. These drugs were identified through the U.S. Food and Drug Administration's National Drug Code database and linked to ChEMBL (a database of bioactive molecules) and the Medical Expenditure Panel Survey (MEPS) Prescribed Medicines Files, yielding analytic samples of 571 for molecular structure, 551 for therapeutic target, and 430 for delivery properties.

Main outcome measures

Annual gross revenues and total prescriptions per product from MEPS and net revenues, incorporating net-to-gross pricing ratios from SSR Health data.

Results

Gross and net revenue rose disproportionately for high-novelty drugs beginning around 2013 across all three novelty dimensions. From 2013 to 2019, mean gross revenue tripled for drugs with high molecular structure novelty, more than tripled for those with novel therapeutic targets, and more than doubled for those with novel delivery properties. Over the same period, gross revenue for medium- and low-novelty groups remained relatively stable across all dimensions. The increase in revenue for high-novelty drugs was driven primarily by growth in prescription volumes, as prices rose at similar rates across novelty groups.

Conclusions

Since 2013, US outpatient drug revenues have risen disproportionately for higher novelty products, driven by growth in share of prescriptions. The trend coincides with U.S. Pharmacy Benefit Managers introducing "formulary exclusion" policies, which may have disproportionately reduced the use of drugs with closer substitutes and thus increased the market returns to pharmacologic novelty.

INTRODUCTION

Pharmaceutical innovation has long been a cornerstone of medical progress, providing new treatment options that improve health outcomes and quality of life. Empirical research shows that firms' investment in innovation and the launch of new drugs are strongly shaped by the revenues those products are expected to generate, making financial reward essential to sustaining innovation.[1–6] Yet debate continues over whether these market rewards primarily accrue to truly novel therapies or to "me-too" drugs; follow-on products that closely resemble the chemical structure or pharmacological function of existing treatments within the same therapeutic class.[7–9] Critics argue that me-too drugs often deliver only modest or incremental clinical benefits, yet still command high prices and generate substantial returns, raising concerns about whether financial incentives are aligned with development of truly transformative therapies.[7–9]

Despite the prominence of these concerns, empirical evidence remains limited on whether pharmacologically novel therapies generate greater revenues than less innovative, follow-on products. A small body of research has linked measures of pharmacological novelty to correlates of financial success, such as inclusion on top-selling drug lists, forward patent citations, and stock market reactions to patent grants, with generally positive associations.[10,11] While these findings suggest that novelty may be rewarded in the marketplace, such proxies do not directly capture actual drug-level revenues. One study found a positive alignment between added clinical benefit, as rated by a health technology assessment organization, and revenues, but this analysis was restricted to oncology drugs.[12]

Compounding this issue is the challenge of measuring pharmacological novelty, which is inherently complex and multidimensional. Many studies rely on simplified, one-dimensional metrics such as first-in-class designation,[13] therapeutic target novelty,[11] or molecular structure dissimilarity.[10] While useful, these indicators often fail to capture the broader technological and pharmacological dimensions that constitute meaningful innovation. For example, a new drug may closely resemble existing drugs in molecular structure, but may have

novel delivery-related features, such as improved absorption, distribution, metabolism, and excretion (ADME) characteristics (for example, pitavastatin is a newer statin that undergoes less metabolism than atorvastatin, despite having a somewhat similar molecular structure[14]) that reduce side effects (for example, development of second- and third-generation sulfonylureas compared to first-generation drugs such as tolbutamide[15]), enhance adherence, or expand treatment to broader patient populations.[16,17] Ratings of added clinical benefit can provide complementary insights, but they often depend on subjective interpretation and are not applied uniformly across contexts. In the absence of an objective, multidimensional framework to capture these varied facets of innovation, current approaches risk misclassifying or undervaluing clinically meaningful advances, thereby limiting our ability to assess how different forms of innovation shape revenues across therapeutic markets.

Another complication lies in the distinction between gross and net drug revenues. Manufacturers often provide substantial rebates to pharmacy benefit managers (PBMs) and insurers, which can significantly lower the effective cost of drugs to the healthcare system.[18] As a result, gross revenue based on list prices may substantially overstate true costs, particularly for drugs with high list prices, which tend to receive larger rebates.[19] Yet it remains unclear whether rebate patterns systematically vary by a drug's level of novelty. Adding further complexity, beginning in 2012, PBMs increasingly adopted closed formularies that excluded coverage for certain drugs, including many newly launched therapies,[20] and PBMs have been more likely to exclude drugs with therapeutically similar alternatives. In this context, it is uncertain whether novel drugs experienced faster revenue growth from fewer exclusions, or whether manufacturers of these drugs were compelled to offer larger rebates in order to secure and maintain formulary placement.

This study addresses these gaps by developing a multidimensional measure of pharmacological novelty across three key domains: molecular structure, therapeutic target, and delivery properties. We construct these measures using data from ChEMBL,[21] a comprehensive database of bioactive molecules with drug-like properties, and link them to nationally representative, drug-level revenue data from the Medical Expenditure Panel Survey

(MEPS)[22] to estimate gross pharmaceutical revenue. To account for discounts and rebates, we incorporate pricing and rebate information from SSR Health[23] to generate estimates of rebate-adjusted net revenue over time. Using this integrated dataset, we examine whether trends in drug revenue are disproportionately driven by highly novel therapies or by products that offer only incremental innovation.[7–11]

METHODS

Data Sources

Our analysis focuses on FDA-approved new molecular entities. To construct a comprehensive, drug-level dataset, we integrated multiple nationally representative and domain-specific data sources. Molecular and pharmacological characteristics were obtained from ChEMBL, a curated database of bioactive compounds with drug-like properties.[21] ChEMBL includes annotated information on molecular structures represented by SMILES strings, molecular targets, physicochemical parameters relevant for evaluating delivery properties, and therapeutic classifications based on the Anatomical Therapeutic Chemical (ATC) system.

Regulatory and formulation data, including approval year, and New Drug Application (NDA) identifiers, were extracted from the U.S. Food and Drug Administration's (FDA) National Drug Code (NDC) data files.[24] Current and discontinued (excluded) NDC datasets were used to capture the full history of each product. Drug gross revenue data were obtained from the MEPS Prescribed Medicines Files for the years 2000 through 2019.[22] MEPS provides nationally representative data on outpatient prescription drug use and spending across all payer types, including private insurance, public insurance programs, and out-of-pocket payments. Finally, pricing and rebate data were obtained from SSR Health,[23] a commercial database that provides quarterly estimates of list prices, net prices, and list-to-net price ratios for branded drugs marketed in the United States.

A multi-step linkage procedure was implemented to construct a unified, drug-level analytic file. First, compounds from ChEMBL were matched to FDA NDC records using standardized generic names. The

resulting dataset was then transformed into a longitudinal panel that follows each drug annually, starting from its market entry year, as identified in the FDA NDC files, and extending through the end of its market exclusivity period, based on archived records from the FDA's Orange Book.[25] For drugs without exclusivity information in the Orange Book archive, we inferred the exclusivity period from the year of first generic entry recorded in the FDA NDC files. This longitudinal panel was then linked to MEPS prescription records using NDC codes and calendar year, to calculate gross annual revenue per drug. Because MEPS reports 11-digit NDCs while FDA files use a 10-digit format, these datasets were aligned by inserting the standard leading or trailing zero used in FDA conversion rules.[26] We retained only those drugs that could be successfully linked across all relevant datasets and had at least one year of observed spending in MEPS. Finally, SSR Health data were linked to the ChEMBL-FDA longitudinal file using brand names and calendar year.

Novelty Metrics

A multidimensional measure of drug novelty was developed encompassing three distinct domains: molecular structure, therapeutic target, and delivery properties.

Molecular structure

Molecular structure novelty was quantified using the Tanimoto similarity coefficient,[27] calculated based on SMILES strings obtained from ChEMBL using the R package *rcdk*.[28] This metric compares the molecular fingerprints of two compounds, producing a score between 0 and 1, where 0 indicates no structural similarity and 1 indicates an identical molecular structure. For each drug, we first calculated a baseline structural novelty score at the time of market launch by identifying the maximum Tanimoto similarity between the drug and all previously approved drugs within the same ATC level 2 class. A lower similarity score indicates higher structural novelty at launch.

To capture changes in relative novelty over time, we implemented a dynamic scoring algorithm. In each subsequent year following market entry, the drug's structural novelty was re-evaluated by comparing it to all new drugs launched that year within the same therapeutic class. If the highest similarity score against these new entrants exceeded the baseline similarity score, the drug's novelty score was updated accordingly; otherwise, it remained unchanged. This dynamic approach reflects how the structural distinctiveness of a drug evolves in the context of emerging competitors, and it allows us to capture the erosion or preservation of novelty over the product lifecycle.

Therapeutic target

Therapeutic target novelty was quantified by counting the number of existing drugs in ChEMBL that shared the same biological (pharmacological) target. A lower count indicates higher novelty, reflecting fewer prior drugs with activity against that target. Unlike binary classifications, such as "first-in-target" versus "not first-in-target", which treat all follow-on drugs equally regardless of how crowded the target space has become, a count-based measure allows for gradations in novelty that may have important implications for scientific risk, clinical differentiation, and market dynamics. For instance, a second-in-target drug may expand clinical choice more meaningfully than a fifteenth-in-target drug, but both effects are assumed identical in a binary framework. Similar methods have been used in prior studies examining the relationship between target novelty and financial outcomes.[11]

Because many drugs act on multiple targets, we computed target-specific novelty scores for each annotated target and then averaged them to produce an overall therapeutic target novelty score. To account for changes over time, we implemented a dynamic scoring framework analogous to that used for molecular structure novelty. At the time of market launch, a drug's baseline novelty was calculated based on the number of previously approved drugs sharing the same target. In each subsequent year, we updated the score to reflect newly launched drugs with overlapping biological targets. As additional competitors with the same target

entered the market, the novelty declined, capturing the reduced distinctiveness of the target mechanism in an evolving therapeutic landscape.

Delivery properties

Delivery-related complexity was evaluated using parameters for physicochemical properties commonly associated with oral bioavailability, primarily guided by Lipinski's Rule of Five.[29] The parameters include molecular weight (MW), lipophilicity (logP), the number of hydrogen bond acceptors (HBA), the number of hydrogen bond donors (HBD), and the number of rotatable bonds (RB), a measure of molecular flexibility that is associated with permeability. While Lipinski's rules are typically used to identify upper thresholds beyond which poor absorption is more likely, we note that molecules with lower values for these parameters may also have poorer absorption that might lead to delivery and formulation challenges. One such example is provided by megestrol acetate, an appetite stimulant used for treatment of cachexia. This molecule has HBA of 4 and HBD of 0, but is essentially undeliverable in patients who have not eaten food (the very condition it is intended for). This problem has been overcome by a formulation change.[30] Therefore, we defined upper thresholds using Lipinski's rules and lower bounds based on our own observations of drug-like properties. Specifically, we applied the following conservative bounds to define optimal delivery characteristics: molecular weight between 350 and 500 Da; logP between 1 and 5; HBA between 5 and 10; HBD between 2 and 5; and RB between 5 and 10. For each drug, a score of 0 was assigned if the value for a given property fell within the optimal range, and 1 if it fell outside the range. The final delivery complexity score was calculated as the average of the five binary indicators, ranging from 0 (optimal across all parameters) to 1 (outside optimal range for all parameters). Higher scores indicate greater deviation from ideal delivery profiles and potentially greater ADME complexities.

Combination therapy

For fixed-dose combination drugs, we adapted our novelty scoring approach to account for the presence of multiple active ingredients. If the combination included at least one newly approved molecule, molecular structure and therapeutic target novelty were assessed based on that new component, as it represents the primary

source of innovation relative to existing treatments. In cases where all components were previously marketed individually, the average novelty score was calculated across all ingredients. For delivery properties novelty, the delivery scores of all components were averaged regardless of approval status, given that the formulation requires them to be co-administered as a single therapeutic entity with shared pharmacokinetic considerations.

Outcome Measures

Gross Revenue

Annual gross drug revenue was calculated using the MEPS Prescribed Medicines Files.[22] For each year from 2000 to 2019, we aggregated total prescription drug expenditures across all observations for each drug. To ensure consistency in the unit of analysis, revenue was aggregated at the level of the NDA code, which represents the FDA's unit of product approval.

Net Revenue (after rebates)

To estimate net revenue, defined as gross revenue after subtracting manufacturer rebates and discounts, we employed a two-stage approach that combines observed pricing data with predictive modeling. First, for drug-year observations covered by both MEPS and SSR Health, we used SSR Health's reported quarterly list and net prices to calculate drug-specific net-to-list ratios. These quarterly rebate ratios were averaged to generate annual estimates. The resulting net-to-list ratios were then applied to the corresponding gross revenue figures from MEPS to calculate net revenue for each drug-year.

Second, for drug-years in the MEPS dataset that were not included in the SSR Health data, we imputed net-to-list price ratios by training a predictive model using SSR data. This applied to approximately 75% of drug-years in the analytic sample. Specifically, we implemented a generalized linear model with a fractional logit link[31] and Least Absolute Shrinkage and Selection Operator (LASSO) regularization to select relevant predictors. The model incorporated the three pharmacological novelty dimensions (molecular structure, therapeutic target, and

delivery properties), therapeutic class, approval year, calendar year, as well as polynomial and interaction terms to flexibly capture potential nonlinearities and interactions. LASSO regularization was used to prevent overfitting and to select a parsimonious set of predictors that minimized out-of-sample prediction error, based on cross-validated mean squared error.[32] The final model was then applied to the MEPS-linked dataset to impute net-to-list ratios for drug-year observations not included in SSR Health. Further details of the imputation procedure are provided in Appendix 1.

Prescription volume and cost per prescription

In addition to total revenues, we examined two related outcomes: (1) prescription volume and (2) average cost per prescription. Prescription volume was defined as the total number of prescriptions per drug-year in MEPS. Average gross and net costs per prescription were calculated as gross and net revenues divided by the corresponding prescription counts, adjusted to 2019 USD.

Statistical Analysis

Unadjusted Trend Analysis

We began by generating unadjusted time trends for both average gross and net revenue at the drug-year level. These trends were stratified by drug novelty within each of the three domains: molecular structure, therapeutic target, and delivery properties. For each domain, drug-year observations were grouped into three discrete novelty categories based on tertiles of the corresponding novelty score distribution. For each novelty group, we calculated annual mean revenue and applied locally weighted scatterplot smoothing (LOWESS) to reduce short-term fluctuations and highlight longer-term trends during the 2000–2019 study period.

Adjusted trend analysis

In addition to unadjusted trends in gross revenue, we estimated adjusted trends using generalized linear models (GLMs) at the drug-year level. To estimate trends for gross annual revenue, we employed a Poisson pseudo–

maximum likelihood estimator, which is well suited for predicting skewed, heteroskedastic outcomes such as revenues.[33] Independent variables included novelty group indicators, calendar year fixed-effects, and their interactions to capture time-varying effects of novelty. All models were further adjusted for therapeutic class and cohort year (i.e., year of launch) fixed-effects. From the fitted models, we derived adjusted mean revenue estimates for each novelty group over time and applied LOWESS smoothing to the predicted trends.

To estimate adjusted net revenues, we first predicted the adjusted net-to-gross ratio using a fractional regression model with the same covariates and fixed effects as above. The fitted ratio was multiplied by the adjusted gross revenue to obtain adjusted net revenue for each novelty group and year. LOWESS smoothing was then applied to the predicted trends.

Analysis of number of prescriptions and costs per prescription

We conducted additional regression analyses to examine adjusted trends in the average number of prescriptions per drug and the average cost per prescription, stratified by novelty group. These estimates help assess whether revenue growth within each group is primarily driven by increased utilization, rising per-prescription costs (as a proxy for price), or a combination of both, and how these drivers differ across novelty categories.

Sensitivity Analyses

We conducted several sensitivity analyses to evaluate the robustness of our findings. First, we examined trends in median revenue to reduce the influence of outliers. Second, we extended the analysis beyond the defined exclusivity periods. Third, we tested the sensitivity of the delivery properties novelty measure to threshold selection by lowering the lower bound of each delivery properties parameter by 25% while keeping the upper bound fixed at the standard Lipinski cutoffs. Fourth, we redefined molecular structure and therapeutic target novelty using each drug's baseline score at the time of approval, rather than a dynamically updated score that evolves over time. Fifth, given prior studies documenting a sharp rise in drug spending around 2013, driven in

part by the introduction of high-cost therapies such as those for hepatitis C,[34–36] we restricted the sample to drugs launched prior to 2013 to assess whether revenue trends among earlier-launch cohorts followed a similar pattern. Finally, we restricted the analysis to drug-years observed in both MEPS and SSR and used the original (non-imputed) net-to-list ratio to calculate net revenue. Due to the limited sample size in this analysis, we dichotomized novelty into high and low groups based on the median score, rather than using three groups.

All statistical analyses were conducted using Stata SE 18.5. Data linkage and cleaning were performed using both RStudio (version 2025.05.1+513) and Stata SE 18.5.

Results

The final sample sizes for drugs with available novelty metrics that were successfully matched to MEPS prescription records from 2000 to 2019 are as follows: molecular structure (n = 573), therapeutic target (n = 551), and delivery properties (n = 430). Appendix 2 presents the flowchart of data linkage and sample selection. In the full sample of drug-years, the mean molecular structure novelty score was 0.70 (SD: 0.29), the mean therapeutic target novelty score was 14 (SD: 17), and the mean delivery properties novelty score was 0.55 (SD: 0.22). 91% of the included drugs were launched before 2011. The most common ATC-2 class was antibacterials for systemic use (7%), followed by ophthalmologicals (5%). Appendix 3 provides additional details stratified by novelty groups.

Figure 1 presents temporal trends in market share, measured as the proportion of total prescription volume, across novelty groups. The market share of low-novelty drugs, defined by molecular structure and therapeutic targets, increased through the early 2010s, whereas their share based on delivery properties remained relatively stable in the earlier years. During this period, high-novelty drugs declined across all three dimensions. In more recent years, however, high-novelty drugs rebounded, with market share rising rapidly across all dimensions.

Figure 2 presents smoothed trends in mean annual gross and net revenue per drug, stratified by novelty group across four dimensions. Beginning around 2013, gross revenue diverged sharply across all dimensions, with high-novelty drugs exhibiting substantially faster growth than their medium- and low-novelty counterparts. For molecular structure novelty, mean gross revenue for high-novelty drugs rose steeply after 2012, from approximately $500 million to nearly $1.5 billion by 2019, while revenue for medium- and low-novelty drugs remained relatively stable. A similar pattern emerged for therapeutic target novelty, where gross revenue for high-novelty drugs increased from about $500 million to more than $1.5 billion, far outpacing the other groups. In the delivery properties dimension, gross revenue for high-novelty drugs grew from roughly $500 million to nearly $1.2 billion, whereas medium- and low-novelty drugs experienced only modest increases.

Although net-to-list price ratios remain similar across novelty groups and decline at a comparable pace over time (Appendix 4), high-novelty drugs are associated with larger absolute rebate amounts (Appendix 5). This reduces, but does not eliminate, the revenue gap between high- and low-novelty drugs. The divergence in net revenues beginning in 2013 largely persists (Figure 2).

Figure 3 presents adjusted revenue trends predicted from the regression models, controlling for therapeutic class and cohort fixed effects. Overall, the adjusted results align with the unadjusted trends in Figure 1, implying that the trends are not driven by changes in the prevalence of certain drug classes or by specific approval cohorts of drugs. Therapeutic target novelty appears particularly influential relative to the other two dimensions (Appendix 6).

Figure 4 presents adjusted mean numbers of prescriptions per drug stratified by novelty group, while Appendix 7 examines trends in adjusted mean cost per prescription. We find no clear divergence in mean cost per prescription across novelty groups beginning in 2013; for example, high- and low-novelty drugs defined by molecular structure show similar increases. In contrast, a marked divergence emerges in utilization: the mean number of prescriptions per drug rises substantially for high-novelty drugs but declines for those with low

novelty. These findings suggest that the divergence in overall spending since 2013 is more likely driven by differences in utilization than by changes in per-prescription pricing.

Appendix Figures 7–13 present results from five sensitivity analyses. First, trends in median revenue closely mirror those in mean revenue, though with smaller magnitudes, indicating that while extreme values affect the mean, they do not drive the overall divergence pattern. Second, when both exclusivity and post-exclusivity periods are included, revenue trends remain consistent with the main analysis. Third, for the delivery properties score based on Lipinski parameters with 25% lower bounds, the divergence between high- and medium-novelty drugs is attenuated, though a large gap compared to low-novelty drugs persists. Fourth, classifying molecular structure and therapeutic target novelty based on baseline scores at launch, rather than dynamic updates, yields results similar to the primary analysis. Fifth, restricting the sample to drugs launched prior to 2013 produced patterns consistent with the main analysis for molecular structure and therapeutic targets. For delivery properties, the divergence between high- and medium-novelty drugs is attenuated, although a substantial gap with low-novelty drugs remained. Finally, when restricting the sample to drug-years observed in both MEPS and SSR, a similar divergence in revenue between high- and low-novelty drugs is observed, with the gap between these groups appearing wider before 2013 for delivery properties.

Discussion

This study introduces a multidimensional framework for measuring pharmacological novelty and examines trends in pharmaceutical revenue by novelty level, using a uniquely assembled dataset that integrates pharmacological characteristics with real-world economic data. By quantifying innovation across three distinct domains, molecular structure, therapeutic target, and delivery properties, we provide a more nuanced and comprehensive approach to evaluating pharmaceutical innovation novelty. Our findings show a pronounced divergence in spending beginning around 2013, with high-novelty drugs experiencing substantially greater growth than their lower-novelty counterparts. Notably, this pattern is consistent across all three novelty

dimensions, indicating that multiple forms of innovation contribute meaningfully to revenue dynamics. The results remain robust in regression models adjusting for therapeutic class and launch cohort, and are further supported by multiple sensitivity analyses.

A key insight from our analysis is the temporal shift in the relationship between pharmacological novelty and revenue. Specifically, the revenue gap between more and less novel drugs has widened substantially since 2013. While prior studies have documented a sharp increase in per capita pharmaceutical spending beginning around 2013, often attributed to the introduction of high-cost therapies such as those for hepatitis C[34–36], they did not examine spending at the drug level or account for differences in drug novelty. Our findings suggest that high-novelty drugs were indeed major contributors to revenue growth after 2013. However, this trend cannot be solely attributed to a small number of newly launched products; we also observe increasing revenue among high-novelty drugs introduced prior to 2013. Notably, median-based analyses, designed to reduce the influence of extreme outliers, reveal similar patterns. Taken together, these results suggest that broader systemic or policy-level shifts, rather than a handful of high-cost drugs alone, may have accelerated revenue growth among more pharmacologically innovative therapies.

One potential explanation relates to the growing use of closed PBM formularies. Beginning in 2012, PBMs increasingly implemented closed formularies that excluded coverage for certain drugs, including many newly launched therapies.[20] The number of excluded drugs has continued to rise over time.[37] Prior research suggests that formulary exclusions often target “me-too” drugs; products with limited differentiation and multiple therapeutic substitutes.[20] In contrast, more innovative drugs may have benefited from preferential formulary inclusion, thereby bolstering their utilization and revenue trajectories. Our findings of increased use of high-novelty drugs after 2012 are consistent with this explanation. Future research should further examine how formulary design and coverage policies interact with multidimensional drug novelty to shape utilization patterns, revenue outcomes, and financial incentives for novel innovation.

The pattern persists for net revenue even after accounting for rebates, although high-novelty drugs are associated with larger total rebate amounts, which partially attenuate the observed gaps across novelty groups. This finding, when considered alongside changing formulary dynamics, suggests a potential mechanism: higher rebates may facilitate favorable formulary inclusion for innovative drugs, enabling them to achieve broader market access.[20,38] As a result, high-novelty drugs may reduce their effective prices through rebates while simultaneously driving increased utilization, ultimately preserving a net revenue advantage over less novel therapies. These findings imply the dual influence of clinical differentiation and rebate strategies in shaping the economic trajectories of pharmaceutical products.

A distinctive strength of this study is its interdisciplinary approach, combining health economics and pharmacological sciences. This collaboration enabled the construction of a unified dataset that links detailed pharmacological characteristics with real-world revenue trajectories, an integration that would not be possible using any single data source. Another strength is the multidimensional nature of our pharmacological novelty framework, which allows a more comprehensive assessment of how different facets of innovation relate to market outcomes. Although all three dimensions show meaningful divergence, therapeutic target novelty exhibits the most pronounced separation, suggesting that differentiation at the mechanism-of-action level may be especially influential in shaping market performance. Together, these findings underscore the importance of adopting a multidimensional perspective when evaluating pharmaceutical innovation and its economic consequences.

Alongside its strengths, the study is subject to several limitations. First, we do not evaluate clinical outcomes or health benefits directly and thus cannot assess whether higher spending aligns with greater therapeutic value. Measuring clinical benefit objectively across a broad range of therapeutic areas remains a challenge. Future research could address this gap within specific disease areas; for example, by linking novelty measures to overall survival improvements in oncology.[39] Second, our analysis excludes biologics, as molecular structure novelty remains difficult to apply to these products; alternative metrics would be needed to assess novelty for

biologics. Third, the MEPS data may underrepresent low-prevalence conditions and recently launched drugs, which could limit generalizability for certain therapeutic areas. Lastly, although our framework captures multiple dimensions of novelty, it does not encompass all relevant facets of pharmacological innovation, such as toxicity profiles. Future research could enhance the framework by incorporating these additional elements, contingent upon the availability of reliable data.

This study provides empirical evidence that pharmaceutical spending has become increasingly concentrated on products exhibiting higher degrees of pharmacological novelty across multiple dimensions. This pattern has intensified over the past decade, indicating that both market dynamics and policy mechanisms may be evolving to favor more differentiated and innovative therapies. These findings underscore the growing economic importance of multidimensional innovation and highlight the need for continued research into how clinical value, reimbursement policies, and formulary decisions collectively shape the financial landscape of drug development and adoption.

References

1. Nordhaus WD. An Economic Theory of Technological Change. *The American Economic Review*. 1969;59(2):18-28.
2. Acemoglu D, Linn J. Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry*. *The Quarterly Journal of Economics*. 2004;119(3):1049-1090. doi:10.1162/0033553041502144
3. Yin W. Market incentives and pharmaceutical innovation. *J Health Econ*. 2008;27(4):1060-1077. doi:10.1016/j.jhealeco.2008.01.002
4. Blume-Kohout ME, Sood N. Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development. *J Public Econ*. 2013;97:327-336. doi:10.1016/j.jpubeco.2012.10.003
5. Filson D. A Markov-perfect equilibrium model of the impacts of price controls on the performance of the pharmaceutical industry. *The RAND Journal of Economics*. 2012;43(1):110-138. doi:10.1111/j.1756-2171.2012.00159.x
6. Dubois P, de Mouzon O, Scott-Morton F, Seabright P. Market size and pharmaceutical innovation. *The RAND Journal of Economics*. 2015;46(4):844-871. doi:10.1111/1756-2171.12113
7. Guzick DS, Adashi EY. The ‘Invisible Hand’ Doesn’t Work for Prescription Drugs. *The American Journal of Medicine*. 2023;136(4):333-335. doi:10.1016/j.amjmed.2022.11.022

8. Eichler HG, Hurts H, Broich K, Rasi G. Drug Regulation and Pricing--Can Regulators Influence Affordability? *N Engl J Med*. 2016;374(19):1807-1809. doi:10.1056/NEJMp1601294

9. Gastala NM, Wingrove P, Gaglioti A, Petterson S, Bazemore A. Medicare Part D: Patients Bear The Cost Of "Me Too" Brand-Name Drugs. *Health Aff (Millwood)*. 2016;35(7):1237-1240. doi:10.1377/hlthaff.2016.0146

10. Krieger J, Li D, Papanikolaou D. Missing Novelty in Drug Development. *The Review of Financial Studies*. 2022;35(2):636-679. doi:10.1093/rfs/hhab024

11. Dranove D, Garthwaite C, Hermosilla M. Does consumer demand pull scientifically novel drug innovation? *The RAND Journal of Economics*. 2022;53(3):590-638.

12. Brinkhuis F, Goettsch WG, Mantel-Teeuwisse AK, Bloem LT. Added benefit and revenues of oncology drugs approved by the European Medicines Agency between 1995 and 2020: retrospective cohort study. *BMJ*. 2024;384:e077391. doi:10.1136/bmj-2023-077391

13. Jena AB, Calfee JE, Mansley EC, Philipson TJ. 'Me-Too' Innovation in Pharmaceutical Markets. *Forum Health Econ Policy*. 2009;12(1):5.

14. Chan P, Shao L, Tomlinson B, Zhang Y, Liu ZM. An evaluation of pitavastatin for the treatment of hypercholesterolemia. *Expert Opin Pharmacother*. 2019;20(1):103-113. doi:10.1080/14656566.2018.1544243

15. Abdelmoneim AS, Hasenbank SE, Seubert JM, Brocks DR, Light PE, Simpson SH. Variations in tissue selectivity amongst insulin secretagogues: a systematic review. *Diabetes Obes Metab*. 2012;14(2):130-138. doi:10.1111/j.1463-1326.2011.01496.x

16. Palmer AK, Jensen MD. Metabolic changes in aging humans: current evidence and therapeutic strategies. *J Clin Invest*. 2022;132(16):e158451. doi:10.1172/JCI158451

17. Baryakova TH, Pogostin BH, Langer R, McHugh KJ. Overcoming barriers to patient adherence: the case for developing innovative drug delivery systems. *Nature Reviews Drug Discovery*. 2023;22(5):387. doi:10.1038/s41573-023-00670-0

18. Mattingly TJ, Hyman DA, Bai G. Pharmacy benefit managers: history, business practices, economics, and policy. In: *JAMA Health Forum*. Vol 4. American Medical Association; 2023:e233804-e233804.

19. Yeung K, Dusetzina SB, Basu A. Association of branded prescription drug rebate size and patient out-of-pocket costs in a nationally representative sample, 2007-2018. *JAMA network open*. 2021;4(6):e2113393-e2113393.

20. Agha L, Kim S, Li D. Insurance design and pharmaceutical innovation. *American Economic Review: Insights*. 2022;4(2):191-208.

21. Gaulton A, Bellis LJ, Bento AP, et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res*. 2012;40(Database issue):D1100-1107. doi:10.1093/nar/gkr777

22. Agency for Healthcare Research and Quality. Medical Expenditure Panel Survey: Prescribed Medicines Files. 2025. https://meps.ahrq.gov/mepsweb/data_stats/download_data_files_results.jsp?cboDataYear=All&cboDataTyp

eY=2%2CHousehold+Event+File&buttonYearandDataType=Search&cboPufNumber=All&SearchTitle=Prescribed+Medicines

23. SSR Health. US Prescription Brand Net Pricing Data and Analysis. 2025. Accessed July 11, 2025. http://www.ssrhealth.com/

24. U.S. Food and Drug Administration. National Drug Code Database Background Information. *FDA*. Published online April 26, 2019. Accessed October 22, 2024. https://www.fda.gov/drugs/development-approval-process-drugs/national-drug-code-database-background-information

25. Williams H. Orange Book patent and exclusivity data - 1985-2016. National Bureau of Economic Research. 2018. Accessed August 7, 2025. https://www.nber.org/research/data/orange-book-patent-and-exclusivity-data-1985-2016

26. Maryland Department of Health. National Drug Code (NDC) Conversion Table. https://health.maryland.gov/phpa/oideor/immun/shared%20documents/handout%203%20-%20ndc%20conversion%20to%2011%20digits.pdf

27. Bajusz D, Rácz A, Héberger K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of cheminformatics*. 2015;7(1):20.

28. Voicu A, Duteanu N, Voicu M, Vlad D, Dumitrascu V. The rcdk and cluster R packages applied to drug candidate selection. *Journal of Cheminformatics*. 2020;12(1):3. doi:10.1186/s13321-019-0405-0

29. Lipinski CA. Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. *Adv Drug Deliv Rev*. 2016;101:34-41. doi:10.1016/j.addr.2016.04.029

30. Deschamps B, Musaji N, Gillespie JA. Food effect on the bioavailability of two distinct formulations of megestrol acetate oral suspension. *Int J Nanomedicine*. 2009;4:185-192. doi:10.2147/ijn.s6308

31. Papke LE, Wooldridge JM. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*. 1996;11(6):619-632. doi:10.1002/(SICI)1099-1255(199611)11:6%3C619::AID-JAE418%3E3.0.CO;2-1

32. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 1996;58(1):267-288.

33. Silva JS, Tenreyro S. The log of gravity. *The Review of Economics and statistics*. Published online 2006:641-658.

34. Congressional Budget Office. Prescription Drugs: Spending, Use, and Prices. January 19, 2022. Accessed August 7, 2025. https://www.cbo.gov/publication/57772

35. Martin AB, Hartman M, Benson J, Catlin A. National Health Spending In 2014: Faster Growth Driven By Coverage Expansion And Prescription Drug Spending. *Health Affairs*. 2016;35(1):150-160. doi:10.1377/hlthaff.2015.1194

36. Commonwealth Fund. Prescription Drug Spending Why Is the U.S. an Outlier? doi:10.26099/v5m7-yf04

37. Xcenda. Skyrocketing growth in PBM formulary exclusions continues to raise concerns about patient access. May 24, 2022. Accessed August 7, 2025. https://www.xcenda.com/insights/skyrocketing-growth-pbm-formulary-exclusions-concerns-patient-access

38. Cohen JP, Khoury CE, Milne CP, Peters SM. Rising Drug Costs Drives the Growth of Pharmacy Benefit Managers Exclusion Lists: Are Exclusion Decisions Value-Based? *Health services research*. 2018;53:2758-2769.

39. Chen AJ, Hu X, Conti RM, Jena AB, Goldman DP. Trends in the Price per Median and Mean Life-Year Gained Among Newly Approved Cancer Therapies 1995 to 2017. *Value Health*. 2019;22(12):1387-1395. doi:10.1016/j.jval.2019.08.005

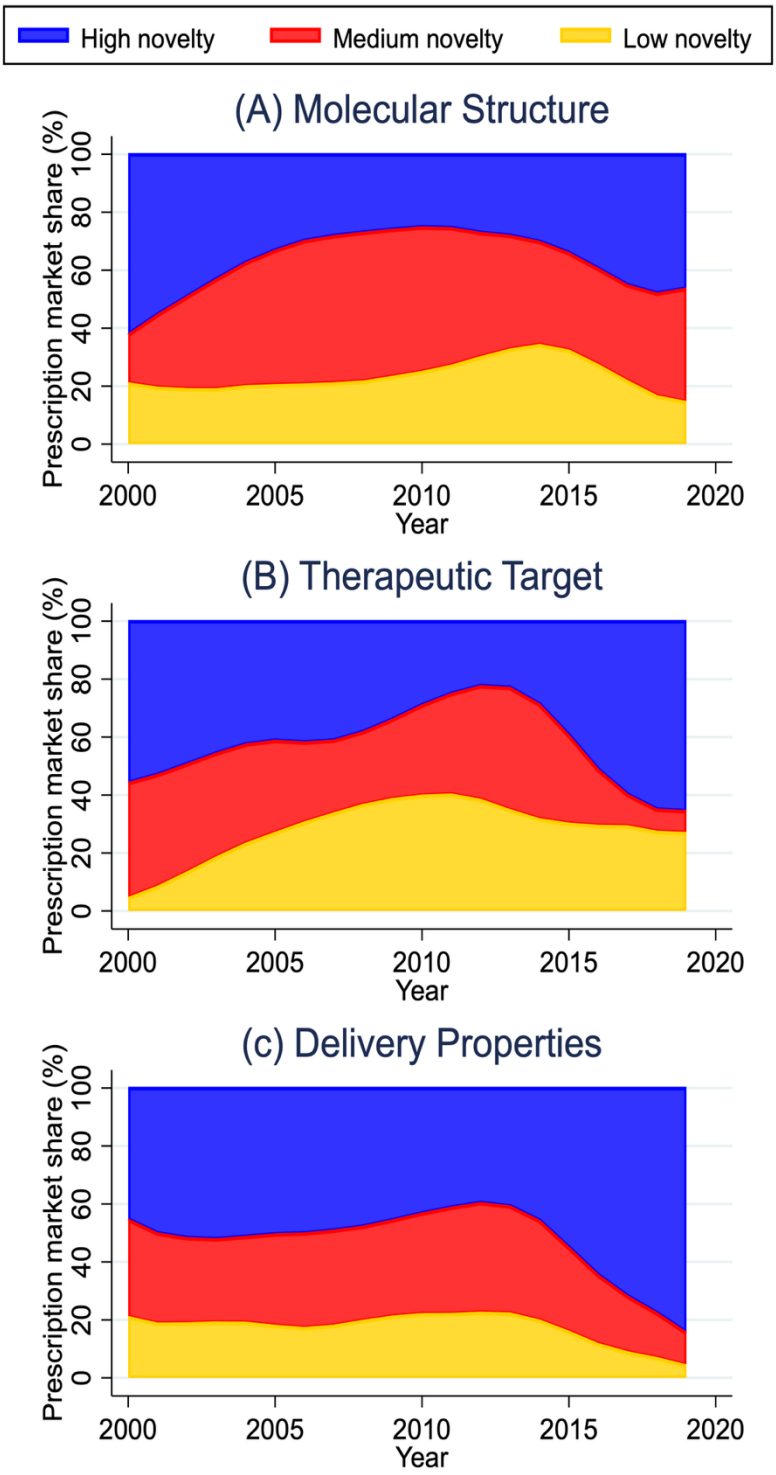


Figure 1. Trends in drug novelty from 2000 to 2019, by novelty dimension

The figure shows the proportion of drugs classified as having high (blue), medium (red), or low (yellow) novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents the smoothed share of drugs on the market in a given year that had not yet reached the end of their exclusivity period and fall into one of the respective novelty categories.

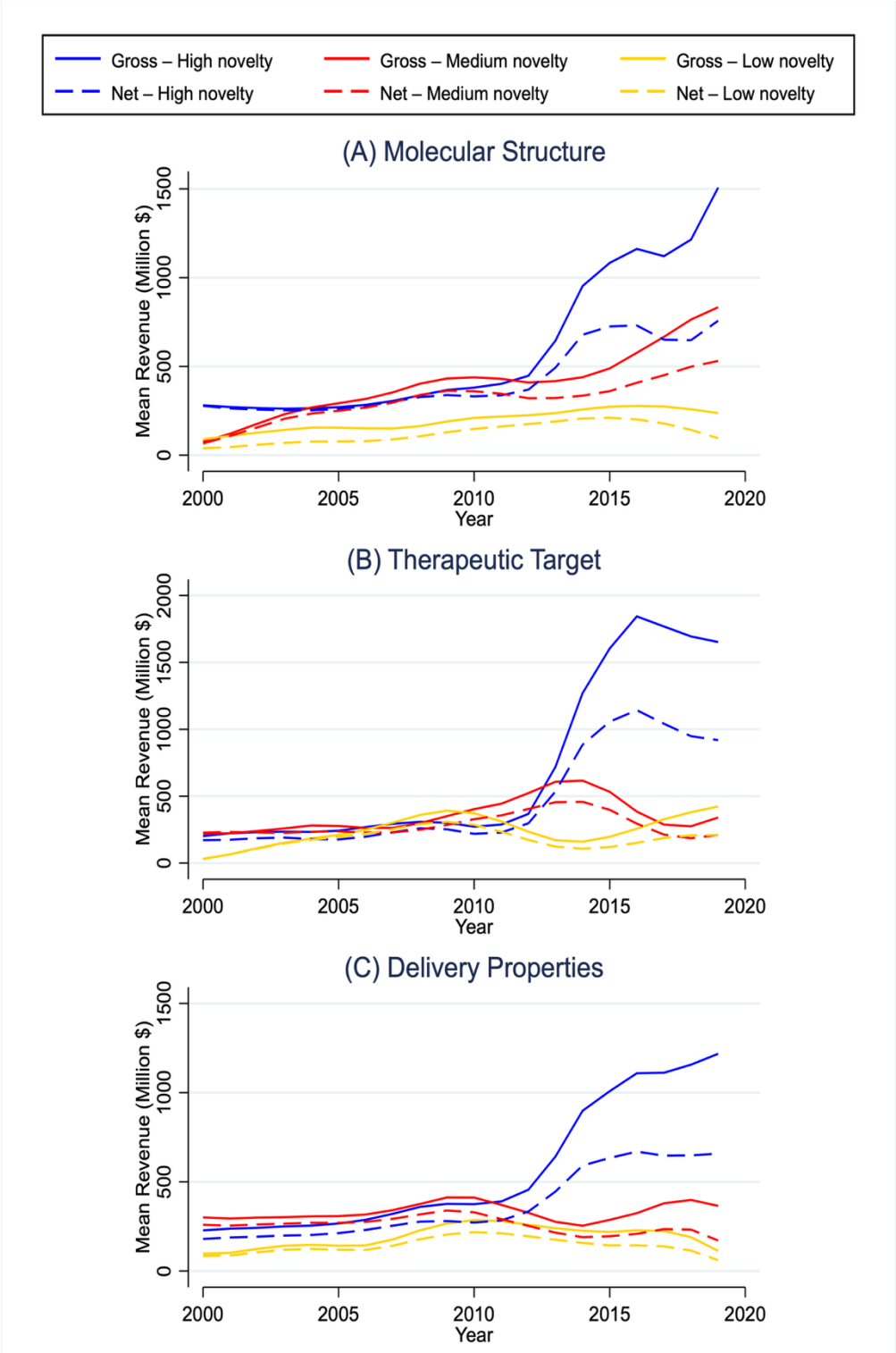


Figure 2. Trends in unadjusted mean gross and net drug revenue by novelty dimension, 2000–2019

The figure shows mean gross and net spending per drug, classified as having high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents the smoothed unadjusted mean revenue (in millions of USD) for drugs on the market in a given year that had not yet reached the end of their exclusivity period and fall into one of the respective novelty categories.

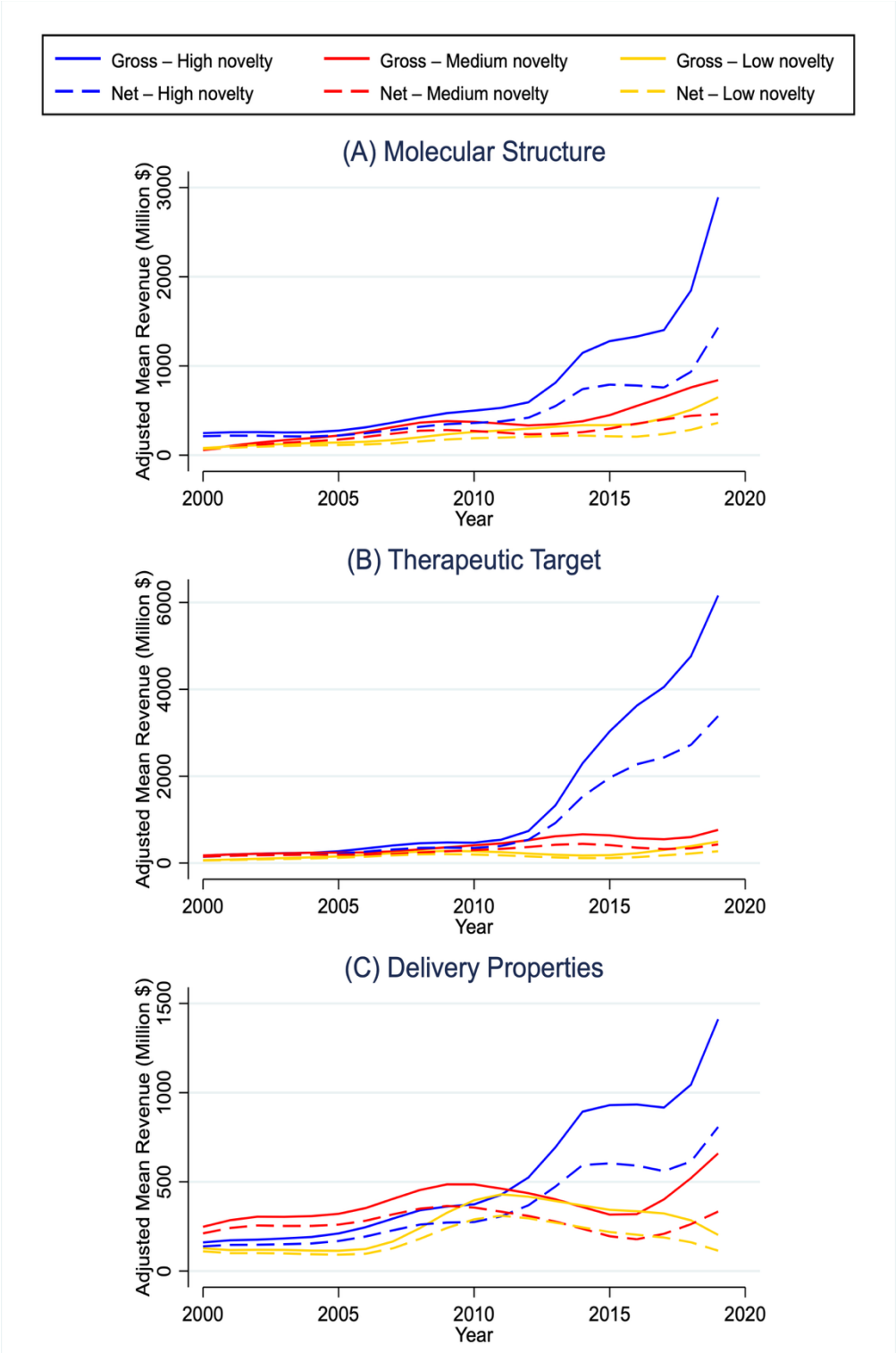


Figure 3. Trends in adjusted mean gross and net drug revenue by novelty dimension, 2000–2019

The figure shows adjusted mean gross and net spending per drug, classified as having high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents the adjusted mean spending (in millions of USD) for drugs on the market in a given year that had not yet reached the end of their exclusivity period and fall into one of the respective novelty or administration categories. Adjusted means are predicted from a regression model controlling for drug class and cohort fixed effects.

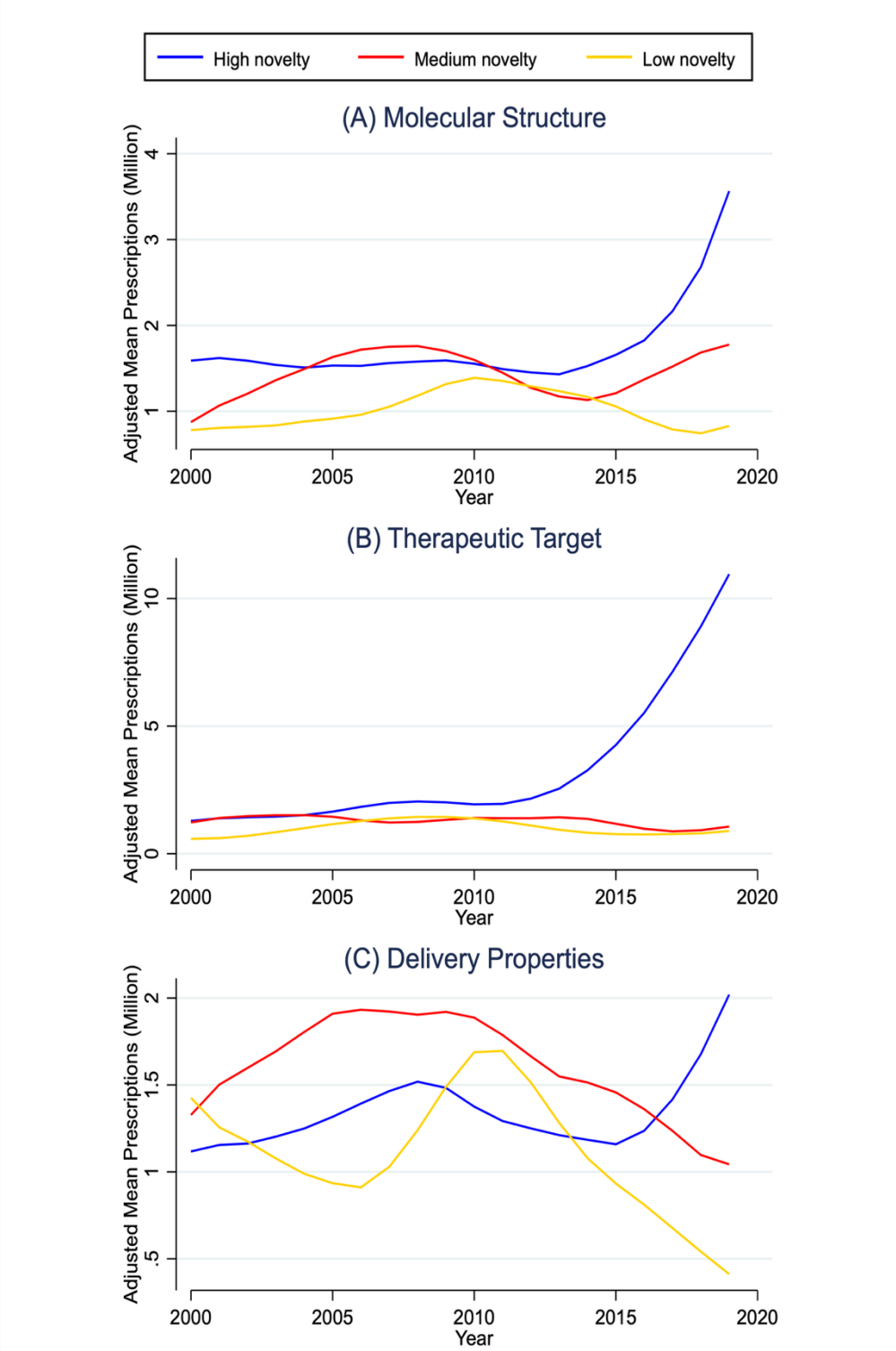


Figure 4. Trends in mean number of prescriptions per drug by novelty dimension, 2000–2019

This figure shows the mean number of prescriptions per drug, classified as high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents adjusted mean number of prescriptions for drugs on the market in a given year that had not yet reached the end of their exclusivity period and fell into one of the respective novelty or administration categories. Adjusted means are predicted from a regression model controlling for drug class and cohort fixed effects.

Supplementary Materials

Appendix 1. Imputation of net-to-list price ratios

Appendix 2. Flowchart of data linkage and sample selection

Appendix 3. Summary characteristics of drugs by novelty dimension and novelty group

Appendix 4. Trends in net-to-list price ratio by novelty dimension, 2000–2019

Appendix 5. Trends in rebates by novelty dimension, 2000–2019

Appendix 6. Trends in Adjusted Mean Gross and Net Drug Revenue by Novelty Dimension, 2000–2019 (Common Y-Axis Scale)

Appendix 7. Trends in adjusted mean costs per prescription by novelty dimension, 2000–2019

Appendix 8. Trends in unadjusted median gross and net drug revenue by novelty dimension, 2000–2019

Appendix 9. Trends in adjusted mean gross and net drug revenue including both exclusivity and post-exclusivity periods, 2000–2019

Appendix 10. Trends in adjusted mean gross and net drug revenue for the delivery properties dimension using lower bounds 25% below the base case, 2000–2019

Appendix 11. Trends in adjusted mean gross and net drug revenue by molecular structure and therapeutic target dimensions using baseline novelty classification, 2000–2019

Appendix 12. Trends in adjusted mean gross and net drug revenue for drugs launched prior to 2013, 2000–2019

Appendix 13. Trends in adjusted mean gross and net drug revenue for drugs observed in both MEPS and SSR, 2000–2019

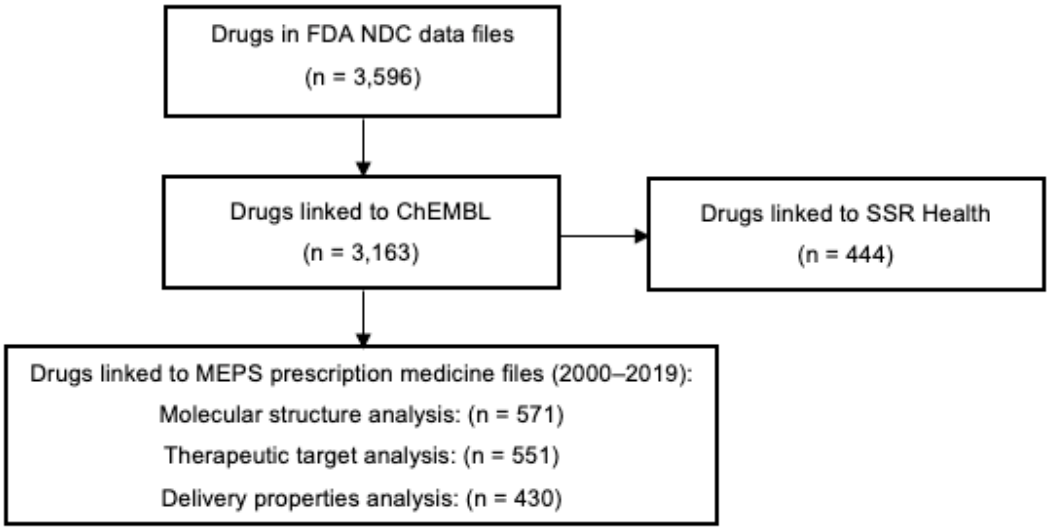
Appendix 1. Imputation of net-to-list price ratios

For drug-years in the Medical Expenditure Panel Survey (MEPS) that were not linked to SSR Health, we imputed net-to-list price ratios using a predictive modeling approach. Approximately 75% of drug-year observations required imputation.

We estimated a generalized linear model with a fractional logit link to account for the bounded nature of the dependent variable (net-to-list price ratio).[31] To prevent overfitting and facilitate predictor selection, we applied Least Absolute Shrinkage and Selection Operator (LASSO) regularization.[32] Model selection was guided by 5-fold cross-validation, minimizing mean squared prediction error. Fractional logit models are well established for proportion data bounded between 0 and 1,[31] and LASSO regularization is widely adopted for high-dimensional predictor selection and improving generalizability in health services and policy research. [32] Together, these methods provide a flexible and robust framework for imputing missing net-to-list ratios while minimizing overfitting.

Candidate predictors included the three drug novelty dimensions (molecular structure, therapeutic target, and delivery properties), therapeutic class, approval year, and calendar year. To flexibly capture nonlinearities and effect modification, we also included polynomial terms (e.g., squared transformations of approval year and calendar year) and the full set of possible interactions among predictors. The final predictor set was selected through the LASSO procedure, yielding a parsimonious specification optimized for out-of-sample predictive performance. This model was then applied to MEPS-linked drug-year observations without SSR Health data to generate imputed net-to-list ratios.

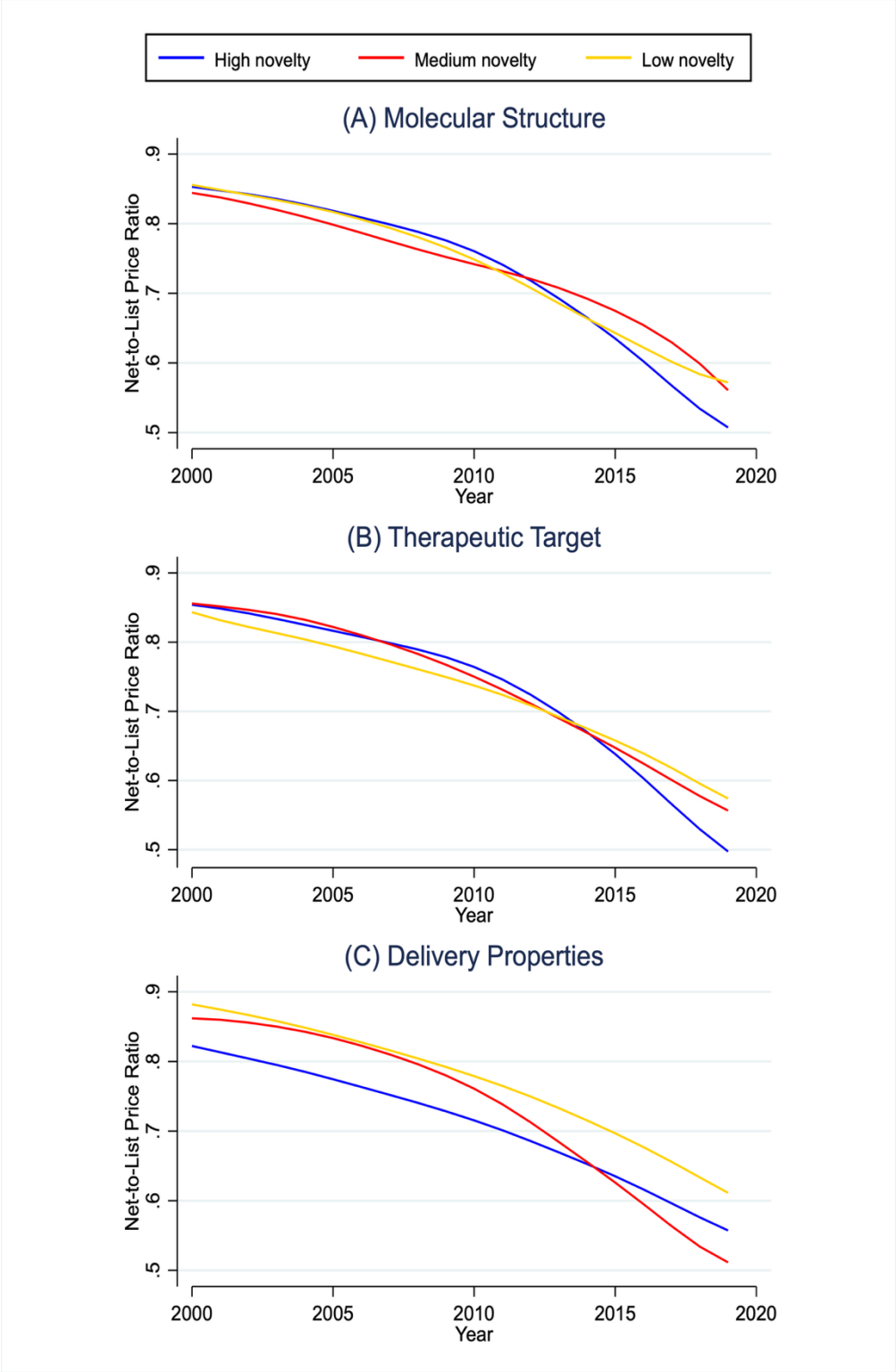
Appendix 2. Flowchart of data linkage and sample selection



Appendix 3. Summary characteristics of drugs by novelty dimension and novelty group

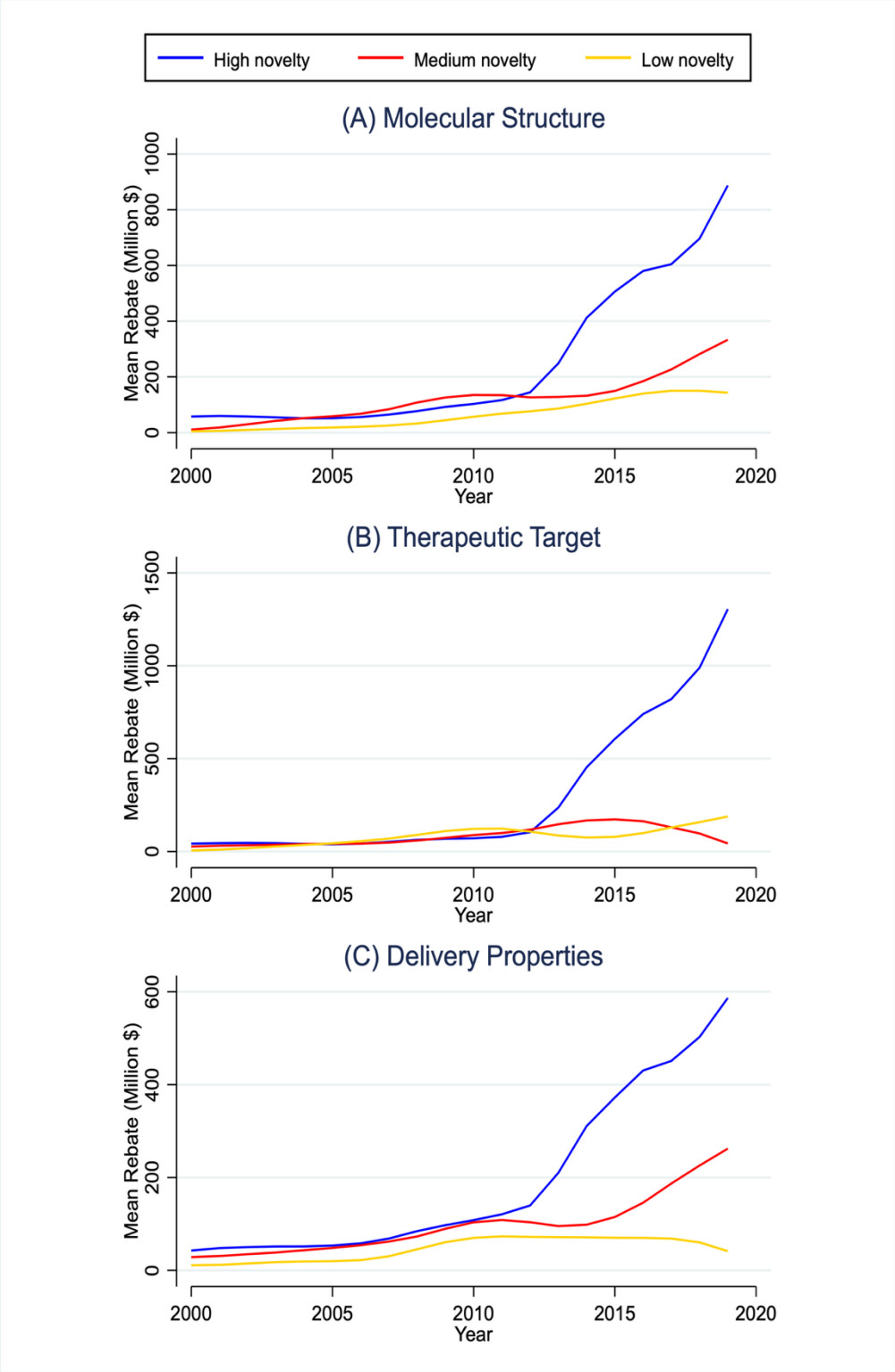
Novelty dimension	Novelty Group	Molecular structure novelty score (mean, range)	Therapeutic target novelty score (mean, range)	Delivery properties novelty score (mean, range)	Pre-2011 launch (%)	Top ATC-2 class (%)	2nd ATC-2cClass (%)
Overall		0.72 (0.00–1.00)	14 (0–159)	0.54 (0.00–1.00)	94%	Antibacterials for systemic use (7%)	Ophthalmologicals (5%)
Molecular structure	High novelty	0.37 (0.00–0.57)	7 (0–66)	0.51 (0.00–1.00)	93%	Antivirals for systemic use (9%)	Psycholeptics (8%)
	Medium novelty	0.73 (0.57–0.99)	14 (0–78)	0.50 (0.00–1.00)	94%	Antibacterials for systemic use (10%)	Sex hormones & modulators (genital system) (7%)
	Low novelty	1.00 (1.00–1.00)	20 (1–159)	0.60 (0.00–1.00)	94%	Sex hormones & modulators (genital system) (8%)	Thyroid therapy (7%)
Therapeutic target	High novelty	0.56 (0.00–1.00)	2 (0–5)	0.51 (0.00–1.00)	93%	Antibacterials for systemic use (9%)	Antibacterials for topical use (6%)
	Medium novelty	0.79 (0.12–1.00)	9 (5–15)	0.55 (0.00–1.00)	96%	Ophthalmologicals (8%)	Antibacterials for systemic use (8%)
	Low novelty	0.85 (0.15–1.00)	32 (15–159)	0.55 (0.00–1.00)	92%	Corticosteroids, dermatological preparations (12%)	Diuretics (9%)
Delivery properties	High novelty	0.68 (0.00–1.00)	14 (0–87)	0.31 (0.00–0.40)	91%	Ophthalmologicals (11%)	Diuretics (7%)
	Medium novelty	0.67 (0.00–1.00)	14 (0–159)	0.59 (0.50–0.60)	93%	Antivirals for systemic use (11%)	Psycholeptics (7%)
	Low novelty	0.80 (0.11–1.00)	17 (0–85)	0.79 (0.70–1.00)	94%	Sex hormones & modulators (genital system) (12%)	Antiepileptics (11%)

Appendix 4. Trends in net-to-list price ratio by novelty dimension, 2000–2019



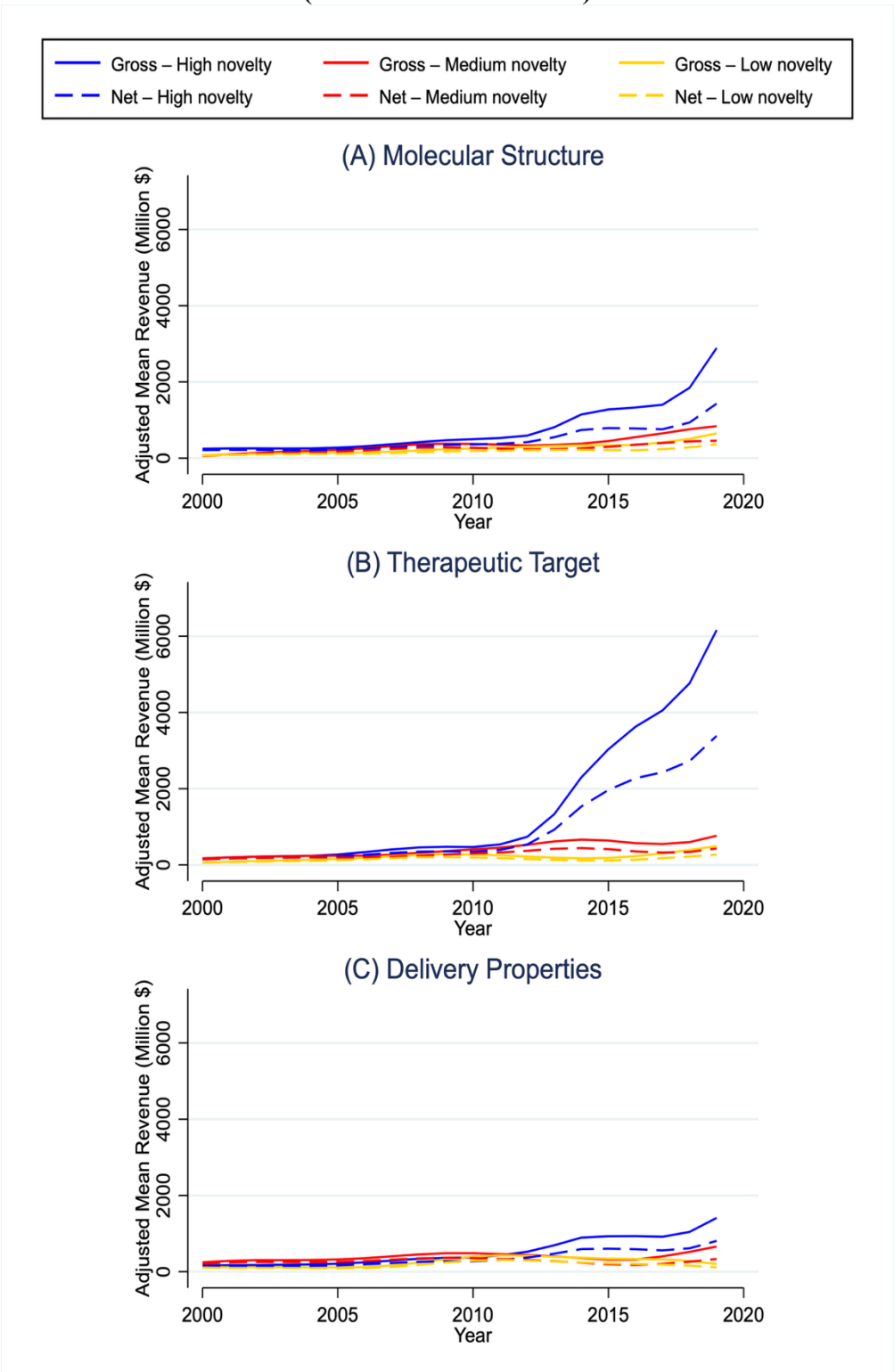
The figure shows unadjusted mean net-to-list price ratio, classified as having high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents the smoothed unadjusted net-to-list price ratio for drugs on the market in a given year that had not yet reached the end of their exclusivity period and fall into one of the respective novelty categories.

Appendix 5. Trends in rebates by novelty dimension, 2000–2019



The figure shows unadjusted mean rebates per drug, classified as having high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents the smoothed unadjusted rebate (in millions of USD) for drugs on the market in a given year that had not yet reached the end of their exclusivity period and fall into one of the respective novelty categories.

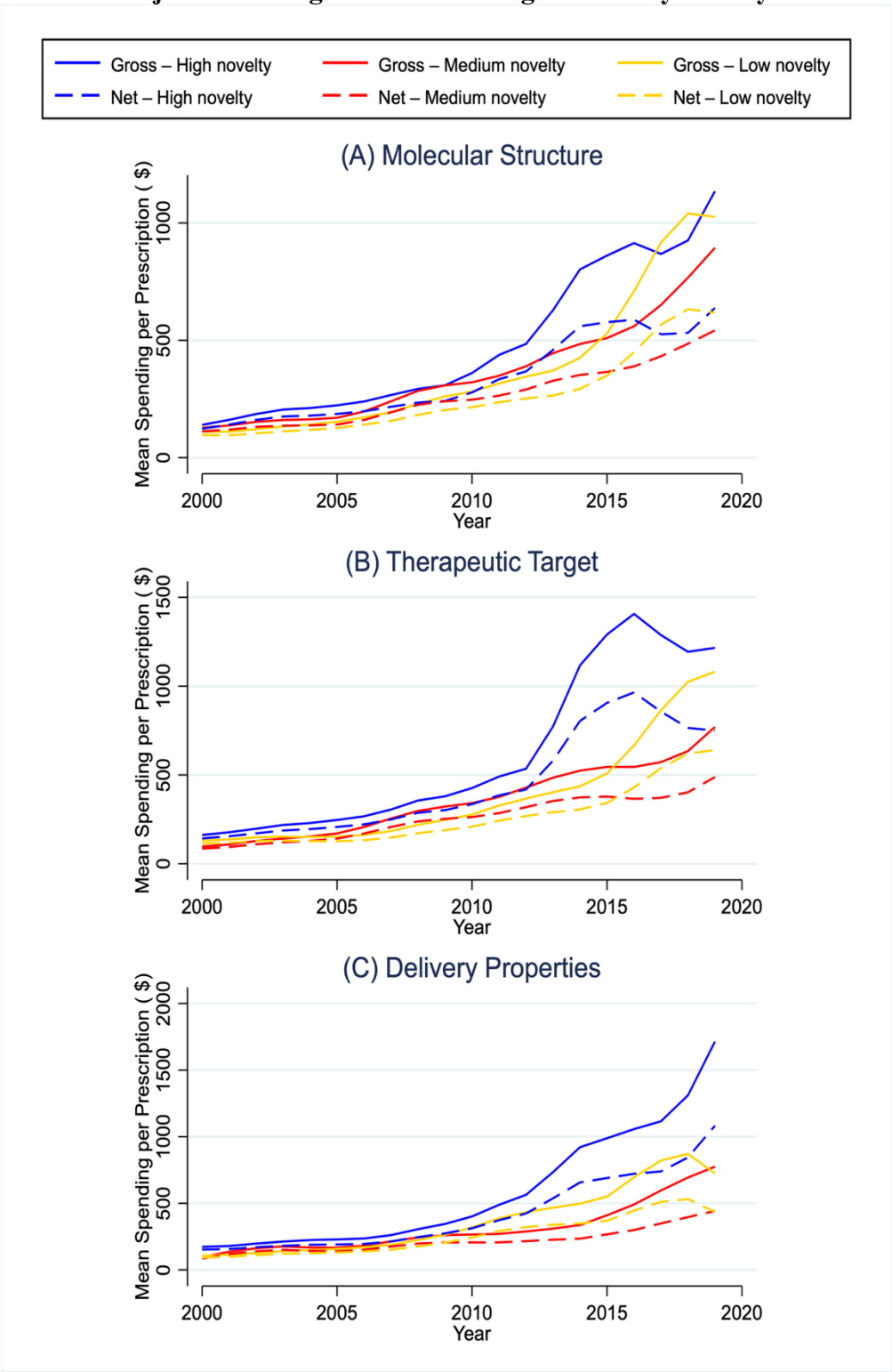
Appendix 6. Trends in Adjusted Mean Gross and Net Drug Revenue by Novelty Dimension, 2000–2019 (Common Y-Axis Scale)



Same as Figure 3, but plotted using a common Y-axis scale for comparability across dimensions. The figure shows trends in adjusted mean gross and net spending per drug (in millions of USD), with drugs classified into high-, medium-, or low-novelty groups for each novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents the adjusted mean spending for drugs on the market in a given year that had not yet reached the end of their exclusivity period and fall within the respective novelty

category. Adjusted means are predicted from regression models controlling for drug class and cohort fixed effects.

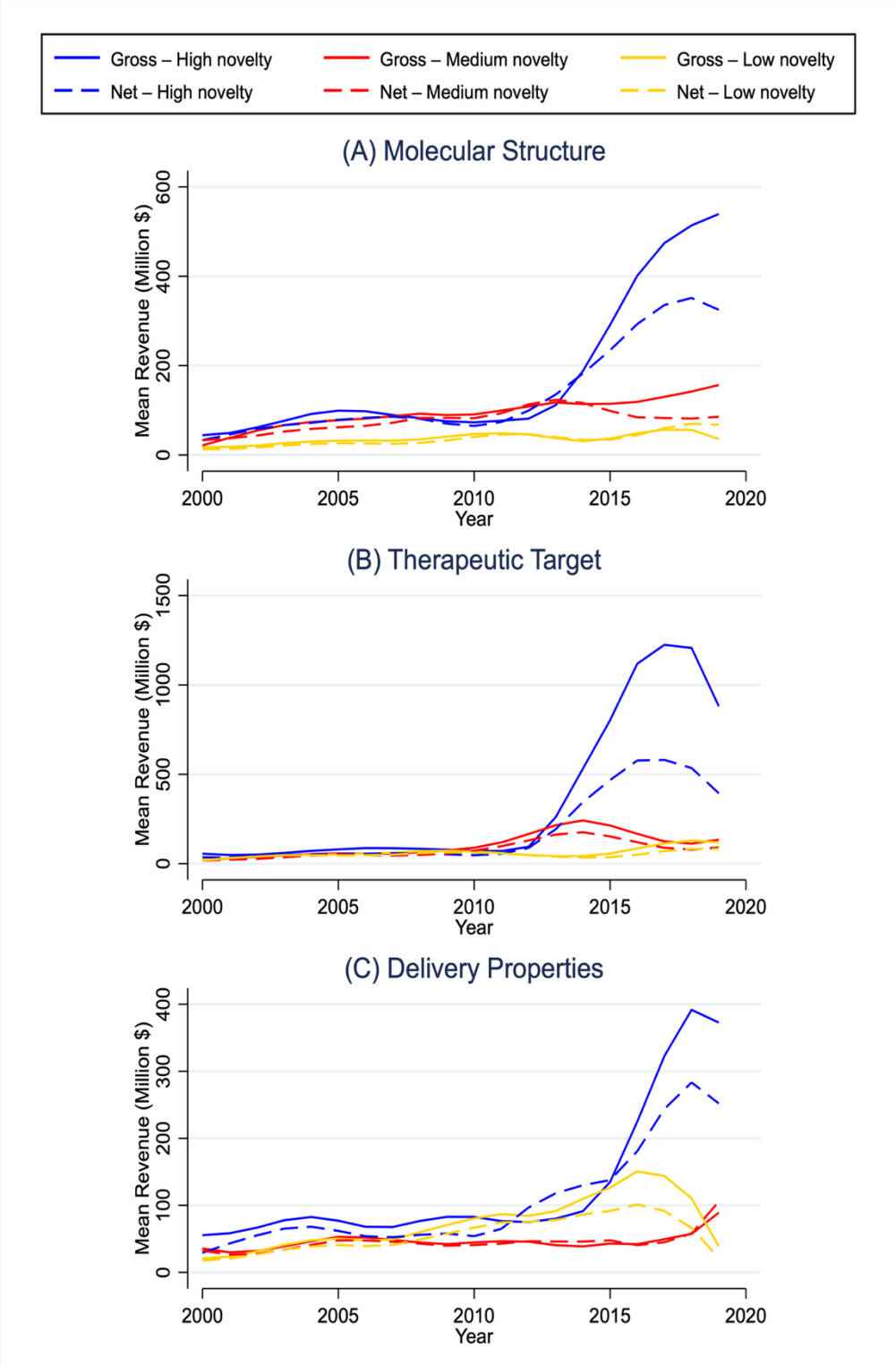
Appendix 7. Trends in adjusted mean gross and net drug revenue by novelty dimension, 2000–2019



This figure shows the mean gross and net costs per prescription, classified as high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents adjusted mean costs per prescription for drugs on the market in a given year that had not yet reached the end of their exclusivity period and fell into one of the respective novelty or administration

categories. Adjusted means are predicted from a regression model controlling for drug class and cohort fixed effects.

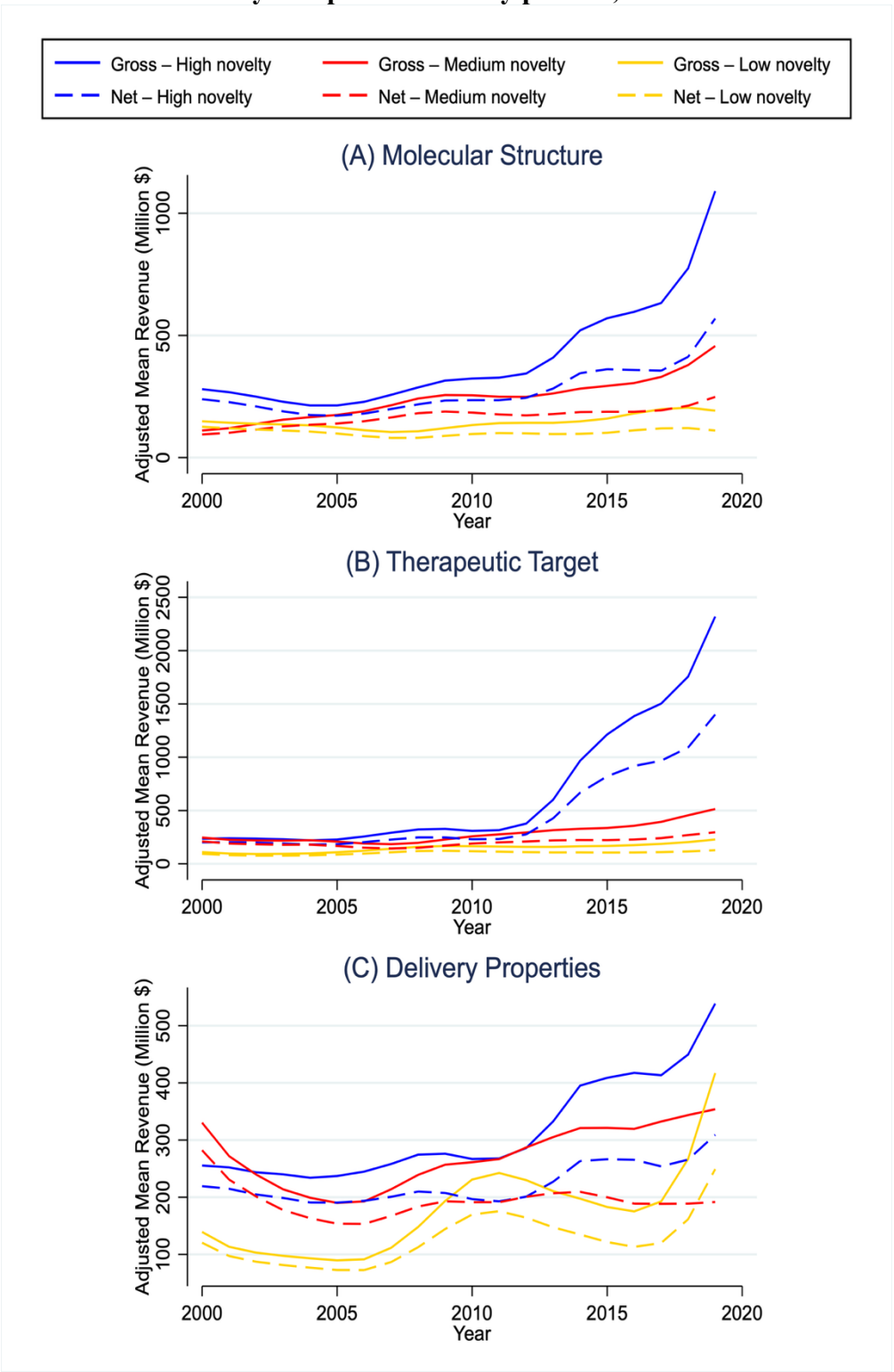
Appendix 8. Trends in unadjusted median gross and net drug revenue by novelty dimension, 2000–2019



The figure shows median gross and net spending per drug, classified as having high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents the smoothed unadjusted median spending (in millions of USD) for drugs on the market in

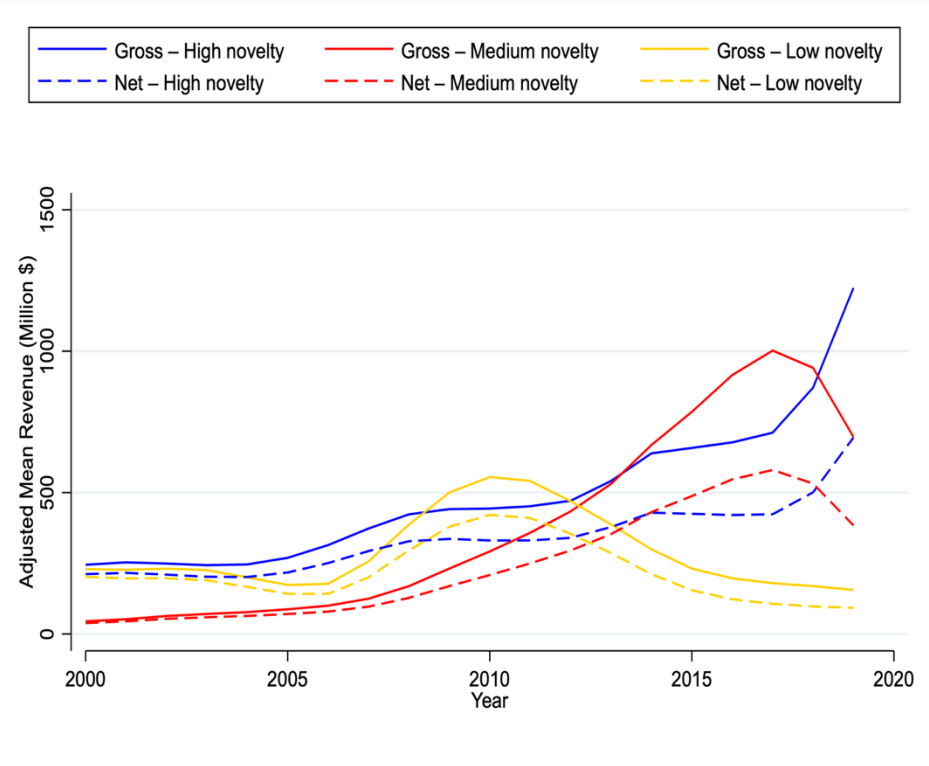
a given year that had not yet reached the end of their exclusivity period and fall into one of the respective novelty categories.

Appendix 9. Trends in adjusted mean gross and net drug revenue by novelty dimension, including both exclusivity and post-exclusivity periods, 2000–2019



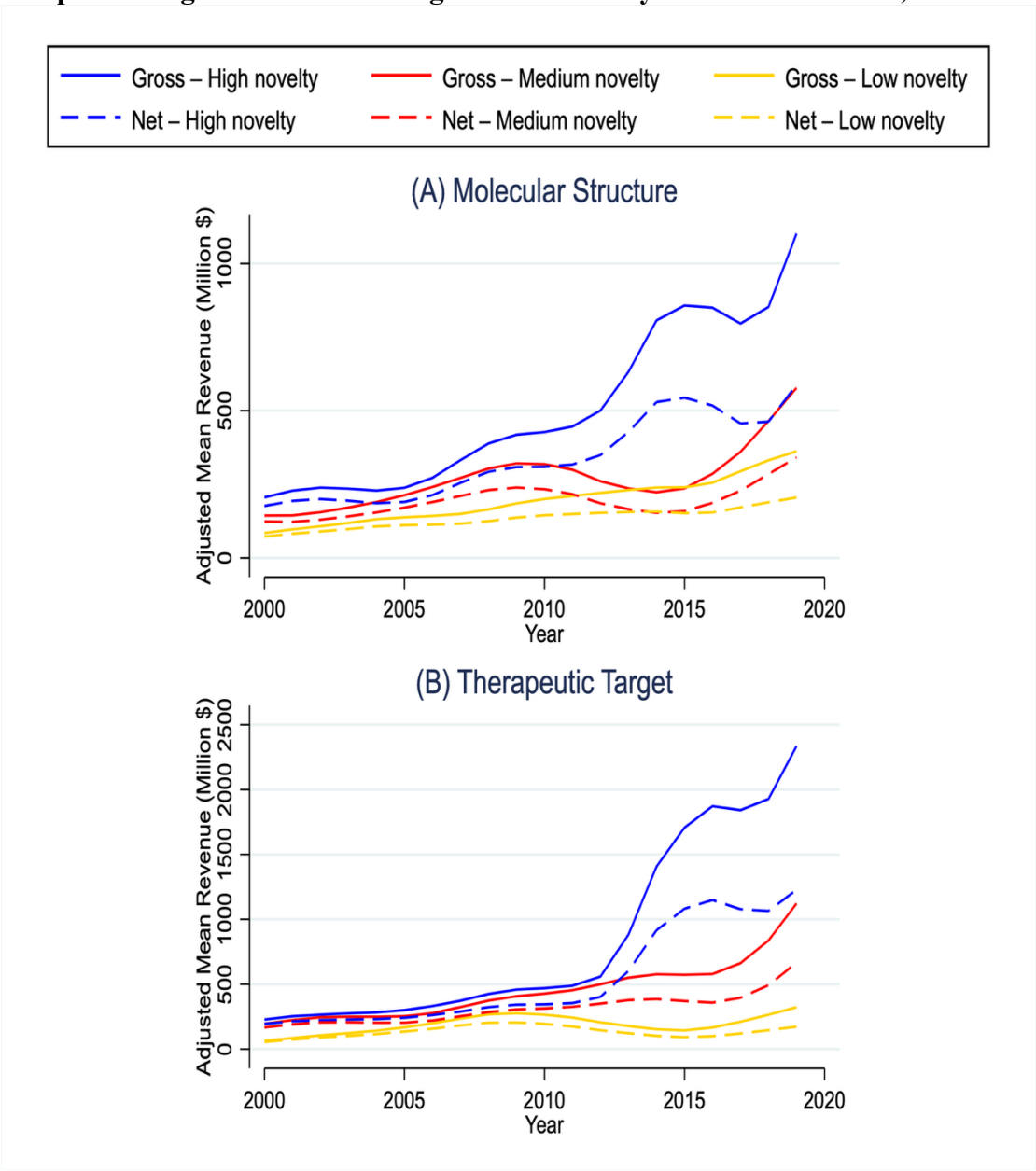
This figure shows the mean gross and net revenue per drug, classified as high, medium, or low novelty, separately by novelty dimension: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Each line represents adjusted mean revenue for drugs on the market in a given year and fell into one of the respective novelty or administration categories. Adjusted means are predicted from a regression model controlling for drug class and cohort fixed effects.

Appendix 10. Trends in adjusted mean gross and net drug revenue for the delivery properties dimension using a lower bound 25% below the base case for score calculation, 2000–2019



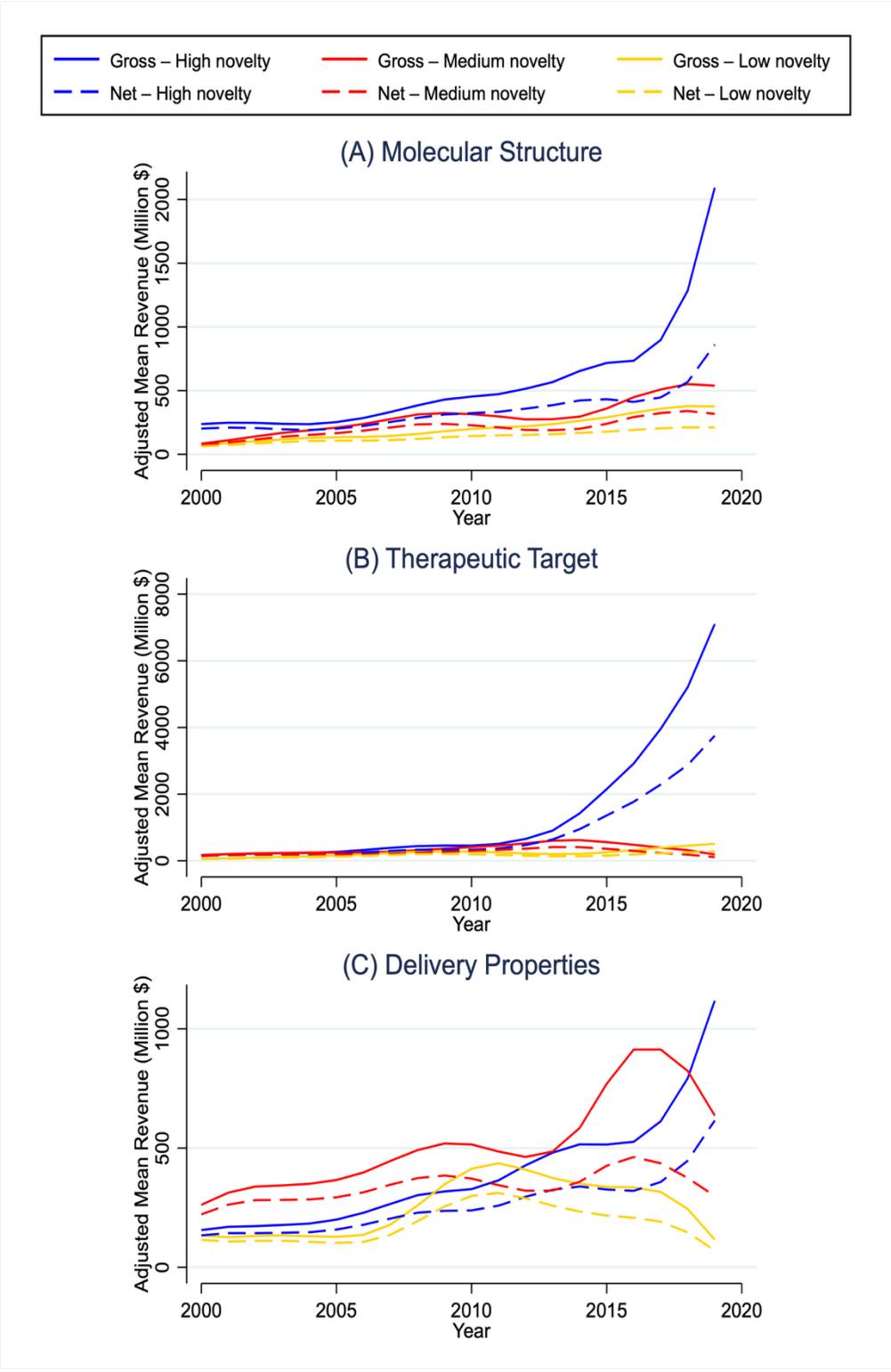
In this sensitivity analysis, novelty scores were recalculated using a lower threshold—25% below the base case lower bound—for score computation. Each line represents adjusted gross and net mean revenue (in millions of USD) for drugs on the market in a given year that had not yet reached the end of their exclusivity period and met the respective novelty criteria. Adjusted means are predicted from a regression model controlling for drug class and cohort fixed effects.

Appendix 11. Trends in adjusted mean gross and net drug revenue by molecular structure and therapeutic target dimensions using baseline novelty score classification, 2000–2019



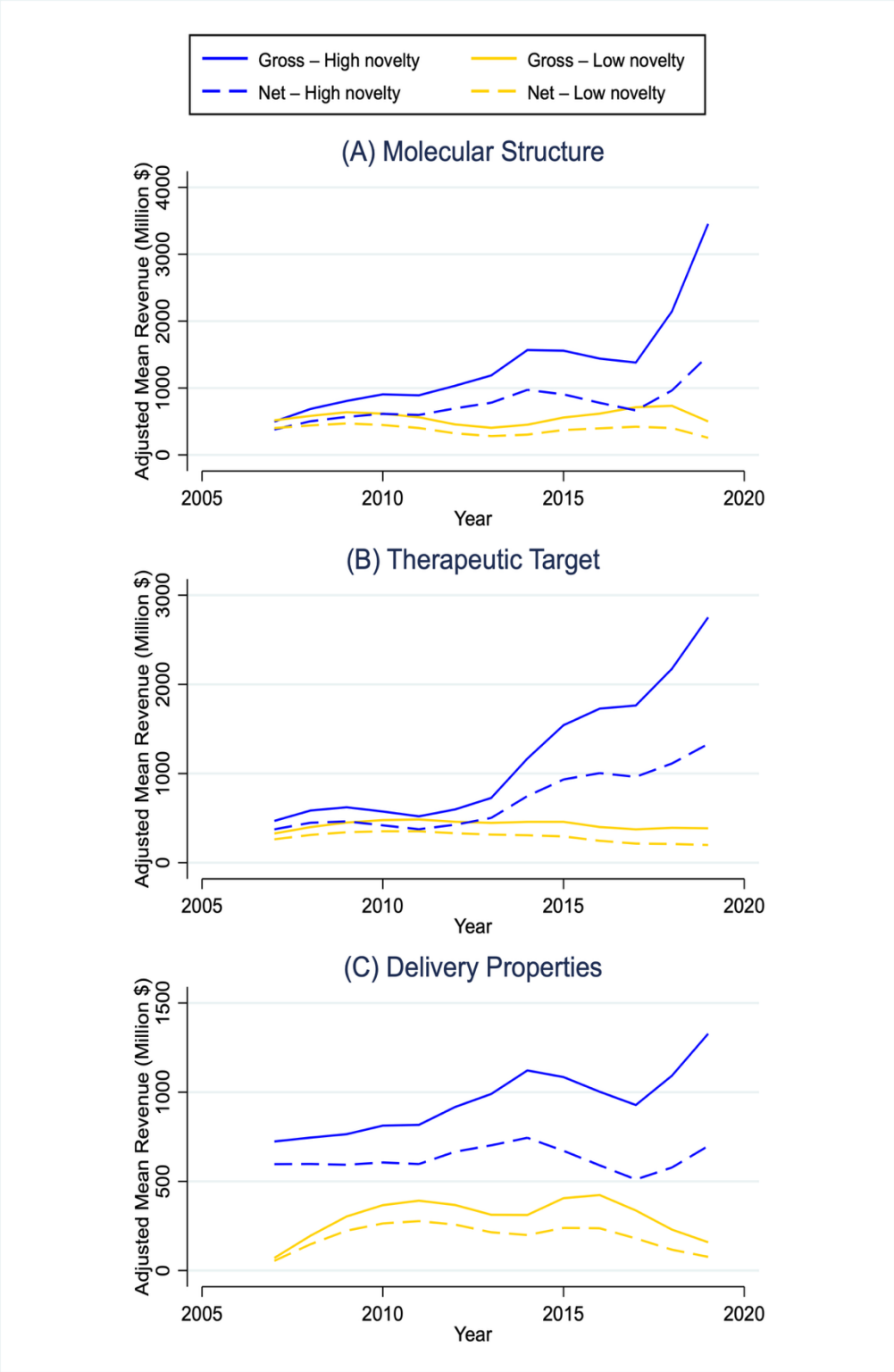
This figure displays unadjusted mean gross and net revenue per drug, stratified by high, medium, and low novelty based on baseline novelty scores for (A) molecular structure and (B) therapeutic target. Each line represents adjusted gross and net mean revenue (in millions of USD) for drugs on the market in a given year that had not yet reached the end of their exclusivity period and met the respective novelty criteria. Adjusted means are predicted from a regression model controlling for drug class and cohort fixed effects.

Appendix 12. Trends in adjusted mean gross and net drug revenue for drugs launched prior to 2013, 2000–2019



This figure shows the unadjusted mean gross and net revenue per drug, categorized by novelty level (high, medium, or low) across three dimensions: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. The sample is restricted to drugs launched before 2013 to examine long-term revenue trends. Each line represents adjusted gross and net mean revenue (in millions of USD) for drugs on the market in a given year that had not yet reached the end of their exclusivity period and met the respective novelty criteria. Adjusted means are predicted from a regression model controlling for drug class and cohort fixed effects.

Appendix 13. Trends in adjusted mean gross and net drug revenue for drugs observed in both MEPS and SSR



This figure presents the adjusted mean gross and net revenue per drug, stratified by novelty level (high vs. low) across three dimensions: (A) molecular structure, (B) therapeutic target, and (C) delivery properties. Net revenue is estimated using the original net-to-gross price ratio. To ensure sufficient sample size, novelty is dichotomized at the median into high and low groups. Each line depicts the adjusted mean revenue (in millions of USD) for drugs on the market in a given year that had not yet lost exclusivity and

satisfied the novelty classification. Adjusted means are derived from regression models controlling for drug class and cohort fixed effects.