EC2C1 Project Coversheet

CANDIDATE NUMBER:	**50196**	**WORD COUNT:**	**2436**
LSE ID:	**202425936**		

Submission instructions

1. You must complete **ALL sections** of this coversheet.
2. You may attach this coversheet to the front of your project, or upload your project and this coversheet as separate files
3. Please upload your project with the coversheet in the dedicated Moodle portal, as well as your **log file**. Please retain a copy of your submission.
4. Please **ensure your name does not appear anywhere in your submission**.
5. Your project should be uploaded either as a .pdf or as a .doc/.docx file (.pdfs are preferred, all else equal)
6. **Name your file** using your candidate number in the following way e.g. 32*784_Project.pdf*. Please name your log file in the same way: e.g. *32784_logfile.log*. Your candidate number is available on LSE for you, and your LSE is on your student card. We need at least one of them to be able to link you to your work
7. You must submit your project, with the completed coversheet, **before the deadline**. Submission after the deadline will result in the application of late penalties as per School guidance here.
8. **You MUST click the 'Submit assignment' button and tick the declaration within Moodle before submission is final.**

Fit to sit /submit

By submitting this project, you declare yourself fit enough to do so. "Fit" in this instance is not only physical or mental health, but also other factors that may affect your academic performance. If you are not fit to sit then please follow the School extension or deferral processes.

Extensions and Deferral

If you are impacted by exceptional circumstances proximate to the project deadline and wish to request an extension, please follow the School extension processes here, copying Steve Pischke (s.pischke@lse.ac.uk) and Tom Glinnan (t.m.glinnan@lse.ac.uk) into your extension request to the Department of Economics.
If your exceptional circumstances are such that a short extension is insufficient, then you are advised to follow the School deferral process here.

Academic Integrity

The policy for the use of AI, such as ChatGPT, is available on Moodle. All uses of AI should be referenced and cited appropriately, according to the guide. Note that your project will be submitted through Turnitin as part of the submission process. By submitting you confirm that you understand School regulations on assessment offences, which can be found here and here. Academic misconduct cases will be pursued in full.

Department of Economics Honour Code

The Department of Economics is committed to the highest standards of academic integrity in every aspect of academic work. All students and staff are expected to uphold these standards both individually and collectively, and to act honourably in the pursuit of learning, teaching and research. Academic misconduct in all its forms contravenes the values of the LSE, as well as those of the wider academic community.

DECLARATION: Please acknowledge the declaration by signing using your candidate number in the box e.g. Candidate 32784.	Candidate 50196. *I solemnly undertake to abide by the Department of Economics Honour Code during this assessment.*

I solemnly undertake to abide by the Department of Economics Honour Code during this assessment.	

Shale well, birds population and human population: A replication and extension of Katovich (2024)

1. Introduction

Katovich (2024) investigates the impact of expansions in shale oil and gas production, as well as wind energy, on wildlife populations and species in the United States from 2000 to 2020. He applies a panel difference-in-differences (DiD) model, using both dummy and continuous variables for the independent variables. Katovich finds that shale production reduces bird populations by 15%, while wind turbines have no significant effect. Neither energy infrastructure significant affects bird species. This replication focuses on the impact of shale well on bird population, which was the only statistically significant result in the original study (Katovich, 2024). In addition, Katovich (2024) treats human population as a mediator between shale well development and bird populations, finding no significant effect and concluding that bird decline is directly caused by shale extraction. However, I argue that human population may confound the relationship between shale wells and bird populations by influencing well siting. Therefore, I include human population as a control to address potential omitted variable bias (OVB).

A key motivation for this modification is to address selection bias. Shale oil and gas development often leads to environmental pollution (Michalski and Ficek, 2016), such as water (Vidic et al., 2013) and air pollution (Institute of Medicine, 2014), which causes public health concerns. Therefore, several states governments implement setback regulations to limit drilling near residential areas, in states Pennsylvania (Miller, 2021), Colorado (Haley et al., 2016), and New Mexico (Mccartney, 2024). In some densely populated regions, such as New York, drilling has even been legally prohibited (Horwitt, 2013). Thus, human population can constrain the shale well's location and are not randomly assigned. Meanwhile, human activity may impact bird populations through urbanization. Therefore, human population is a confounder influencing both shale well and bird populations. By controlling for human population, this model can address the selection bias challenge.

This paper is structured as follows. Section 2 introduces the data and functional form. Section 3 replicates the baseline model to examine the relationship between shale well activity and bird populations, using both binary and continuous outcome specifications. A weighted least squares (WLS) estimation is also applied, weighting on bird counter effort. Section 4 explores the role of human population by including it as a control variable in the continuous specification. In addition, an interaction term between human population and shale well activity is introduced to examine potential heterogeneous effects.

2. Data and measurement

The original paper (Katovich, 2024) uses data on bird populations (y_{ct}) at the circle level (c) from 2000 to 2020 (t) in the US, based on counts from the Christmas Bird Count (CBC). To explore heterogeneity effect across outcomes, the analysis categorizes birds by ecological type, including Grassland, Woodland, Wetland, Other Habitats, Urban, Non-Urban, Non-Migrants, Short and Long Migrants. The unit of observation is a CBC circle by year, resulting in a panel of over 31,000 circle-year observations. As shown in *Table 1*, zero value ratios in the bird count data are lower than 0.2%, thus I use natural log transformation in binary treatment specification on the outcome variables, allowing for an interpretation of coefficients in elasticity terms: $100 \times (e^{\beta} - 1)$. In the continuous dependent variable specification, I replicate Katovich's approach by using the inverse hyperbolic sine (IHS) transformation to keep consistent with the form of treatment variable. To reduce the impact of outliers, the dependent variable is winsorized at the 1st and 99th percentiles.

Data on shale oil and gas wells are collected across the United States for the period 2000–2020. These are treatment variables (T_{ct}), defined at the circle (c) and year (t) level, adopting both binary and continuous forms. In the binary specification, the variable equals 1 if at least one shale well is present in a circle and a given year, and 0 otherwise. In the continuous specification, the total number of shale wells operating in a given circle and year is transformed using the inverse hyperbolic sine (IHS)[1] function, which accommodates zero values and reduces the influence of extreme outliers. I replicate this approach in my analysis.

To explore causal mechanisms of human population, Katovich (2024) uses data on annual county-level human population and take logarithm to control for scale effects.

To consider changes in anthropogenic land use within each CBC circle, Katovich (2024) controls for (i) the number of the bird counter, (ii) weather conditions on the count day, including minimum and maximum temperatures, maximum snowfall, and wind speed, and (iii) the share of land within each circle classified as agricultural, pasture, or developed areas. I replicate these human population data and controls in my analysis.

Table 1. Birds and shale well data summary[2]

	(1) Observed	(2) Mean	(3) S.D.	(4) Min	(5) Max	(6) Zero-value Ratio
Total	31674	27606	214099	12	13100000	0.00%
Grassland	31562	764	1380	0	82319	0.04%
Woodland	31658	2301	30948	1	5008455	0.00%
Wetland	31367	7563	27256	0	1212650	0.12%
Other Habitat	31672	16709	204851	0	12500000	0.00%
Urban	31674	20921	207787	11	12500000	0.00%
Non-Urban	31672	6686	27828	0	1504963	0.00%
Resident	31667	1325	1209	1	20811	0.00%
Short Migration	31674	22552	208808	0	12500000	0.00%
Long Migration	31097	3799	22819	0	1501440	0.11%
Any shale	31674	0.027	0.16	0	1	97.32%
Number of shale wells operating	31674	11.29	166.94	0	5983.11	97.32%
IHS.Number of shale wells operating	31674	0.123	0.85	0	9.39	97.32%

[1] The inverse hyperbolic sine transformation: $IHS(y) = ln(y + \sqrt{y^2 + 1})$

[2] The Stata code used in all tables were developed with assistance from OpenAI's ChatGPT (OpenAI, 2025).

3. Baseline specification

3.1. Binary Treatment Variable

The panel difference-in-differences (DiD) model is defined as followed:

$$y_{ct} = \beta_1 T_{ct} + X'_{ct}\beta_2 + \gamma_c + \delta_t + \epsilon_{ct}$$

where c denotes a CBC circle, t is the year. y_{ct} indicates the number of bird count. T_{ct} is the key variable, which is binary variable of whether there is a shale well in binary specification, or a continuous variable of operating shale well in the countinuous specification.

The vector X'_{ct} are controls including: (1) the number of bird counters. This is uncorrelated covariate and improve estimation precision by reducing standard errors. (2) weather conditions on the count day (minimum and maximum temperatures, maximum snowfall, and wind speed); (3) the shares of agricultural, pasture, and developed land within each circle. These are confounders included to reduce OVB and assumed not to be confounded with bird counts.

The treatment group consists of circles with shale wells, while the control group includes circles without any shale development. Additionally, among the treated units, circles that receive shale wells at a later stage can serve as control units for those that have earlier shale development.

The model includes the circle fixed effect γ_c and time fixed effect δ_t, which account for time-invariant characteristics specific to each circle and possible common shocks across all unit. Including both fixed effects helps reduce potential OVB.

The regressions use the clustered standard errors at the circle level to allow a similar correlation between residuals in different periods within circle, assuming independence across circles.

The coefficient of primary interest is β_1 which is a DiD estimator of impact of the shale well on the bird count outcome. In the binary specification, it measures the change in bird count if there exist a shale well citing in the circle.

A key assumption in panel DiD models is the parallel trends assumption, which requires that treated and control groups would have followed similar trends over time in the absence of treatment. If this assumption holds, any post-treatment divergence from the pre-treatment trend can be interpreted as a treatment effect, implying that the control group serves as an appropriate approximation of counterfactual. To check for the parallel trend assumption, I relax it by adding circle-specific linear trends, allowing each circle to have its own linear time trend in the absence of treatment. Specifically, I add an interaction term between circle dummies and years, omitting the last circle each year as the reference group.

I assume there is no classical measurement error in the treatment variable (number of shale wells). However, the dependent variable (bird counts) may suffer from measurement error, as observers could make mistakes or follow inconsistent counting methods. Such errors enter the regression as part of the residual, reducing the precision of estimates and increasing standard errors, but not introducing bias. To make the casual inference, I assume there are no omitted confounders or selection bias, and the model satisfies the exclusion restriction.

Table 2 presents the results of the binary specification. Column (1) shows that shale well arrival has no statistically significant effect on the total bird population, in contrast to Katovich (2024), who reports a 15% decline. However, columns (3) (13) (15) (19) show statistically significant coefficients, indicating that shale well arrival is associated with reductions in specific bird groups: grassland birds (–19%, significant at the 1% level), resident birds (–12%, 5% level), non-urban birds (–15%, 10% level), and long-distance migratory birds (–18%, 10% level).

To test the parallel trends assumption, columns (2) (4) (6) (8) (10) (12) (14) (16) (18) (20) of *Table 2* show regressions including circle-specific trends. The estimates remain in the same direction and qualitatively similar to the basic specification, except that the effects on resident and long-distance migratory birds become statistically insignificant. These results support the parallel trends assumption and the reasonable of the DiD method.

3.2. Continuous Treatment Variable

In the specification of a continuous treatment variable, the model relies on a stronger parallel trends assumption compared to binary DiD settings. Specifically, in the absence of treatment, areas with different amount of shale wells must have followed similar trends in bird populations. This means that any divergence in trends is attributed to the treatment effect. The continuous specification applies a symmetric IHS transformation to both treatment and outcome variables. As shown by Bellemare and Wichman (2020), the IHS–IHS model approximates a log–log specification when variables are strictly positive and have untransformed means above 10. Since the mean number of shale wells is 11 (*Table 1*), this condition is met, and the treatment coefficient can be interpreted approximately as an elasticity: $100 \times (\mathrm{e}^{\beta} - 1)$.

The coefficient of primary interest β_1 captures the marginal effect of the number of operating shale wells within the circle on the outcome variable, holding other factors constant.

The results of DiD model in continuous form are presented in *Table 3*. Column (1) indicates that a 10% increase in shale wells is associated with a 0.43% decline in the total bird population, consistent with the findings of Katovich (2024). Additionally, a 10% increase in well count is statistically significantly associated with declines across multiple bird groups: grassland birds (–0.44%), woodland (–0.27%), other habitat (–0.43%), urban (–0.39%), non-urban (–0.41%), resident (–0.34%), short-distance migratory (–0.40%), and long-distance migratory (–0.48%).

To check the parallel trends assumption, columns (2) (4) (6) (8) (10) (12) (14) (16) (18) (20) of *Table 3* show regressions that include circle-specific trends. While the estimated coefficients for some groups (e.g., grassland, woodland, and non-urban birds) remain stable and statistically significant, suggesting that parallel trends assumption is held. However, results for other groups (e.g., total, urban, resident, short-distance migrants and long-distance migrants) show either larger magnitudes or lower significance. These differences suggest potential violations of the parallel trends assumption. Therefore, these results should be treated with caution.

3.3. WLS Estimation of Continuous Treatment Variable

Table 4 presents the results from weighted least squares (WLS) regressions, where each observation is weighted by the level of counter effort. Since bird counters tend to visit areas with higher bird numbers or greater ecological significance, these observations may contain more valuable information. Weighting these observations allows the analysis to place greater emphasis on those areas. Column (1) shows that a 10% increase in shale production is associated with a 0.37% decline in total bird number. The overall results are similar to those from the continuous treatment specification in *Table 3*, indicating the robustness of the main findings.

Table 2. Shale well and bird population: binary form results

	Total	Total	Grassland	Grassland	Woodland	Woodland	Wetland	Wetland	Other Habitat	Other Habitat	Urban	Urban	Non-Urban	Non-Urban	Resident	Resident	Short Migration	Short Migration	Long Migration	Long Migration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Any Shale	-0.0864	-0.100	-0.208***	-0.229***	-0.0659	-0.0670	-0.112	-0.0975	-0.0116	-0.0277	-0.0622	-0.0750	-0.162*	-0.186**	-0.128**	-0.0750	-0.0655	-0.104	-0.202*	-0.188
	(0.376)	(0.198)	(0.003)	(0.004)	(0.357)	(0.317)	(0.316)	(0.438)	(0.930)	(0.806)	(0.580)	(0.381)	(0.071)	(0.030)	(0.024)	(0.221)	(0.543)	(0.255)	(0.064)	(0.166)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle-specific trend	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	25937	25898	25845	25804	25929	25890	25643	25602	25935	25896	25937	25898	25936	25897	25933	25894	25936	25897	25437	25388
R-sq	0.846	0.875	0.771	0.804	0.805	0.838	0.877	0.898	0.825	0.858	0.820	0.854	0.855	0.879	0.828	0.871	0.845	0.873	0.813	0.842

Note: p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The numbers in parentheses represent standard errors, adjusted for heteroskedasticity and clustered at the state level.

Table 3. Shale well and bird population: continuous form results

	Total	Total	Grassland	Grassland	Woodland	Woodland	Wetland	Wetland	Other Habitat	Other Habitat	Urban	Urban	Non-Urban	Non-Urban	Resident	Resident	Short Migration	Short Migration	Long Migration	Long Migration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Operating Shale Well	-0.0442***	-0.0898*	-0.0449***	-0.0568**	-0.0271*	-0.0392**	-0.0299	-0.0142	-0.0436*	-0.0827	-0.0400**	-0.0775	-0.0421**	-0.0588**	-0.0343***	-0.0239	-0.0410**	-0.0894*	-0.0496**	-0.0506
	(0.008)	(0.058)	(0.002)	(0.022)	(0.055)	(0.043)	(0.186)	(0.682)	(0.056)	(0.132)	(0.033)	(0.118)	(0.015)	(0.030)	(0.004)	(0.207)	(0.024)	(0.070)	(0.044)	(0.164)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle-specific trend	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	25937	25898	25856	25815	25929	25890	25676	25635	25936	25897	25937	25898	25937	25898	25933	25894	25937	25898	25471	25423
R-sq	0.846	0.875	0.772	0.805	0.805	0.838	0.877	0.898	0.825	0.857	0.820	0.854	0.855	0.879	0.828	0.871	0.845	0.872	0.813	0.842

Note: p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The numbers in parentheses represent standard errors, adjusted for heteroskedasticity and clustered at the state level.

Table 4. Shale well and bird population: weighted continuous form results

	Total	Grassland	Woodland	Wetland	Other Habitat	Urban	Non-Urban	Resident	Short Migration	Long Migration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Operating Shale Well	-0.0376***	-0.0392***	-0.0250*	-0.0204	-0.0386**	-0.0350**	-0.0361**	-0.0326***	-0.0344**	-0.0522**
	(0.006)	(0.007)	(0.067)	(0.229)	(0.035)	(0.022)	(0.029)	(0.007)	(0.021)	(0.026)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle-specific trends	No	No	No	No	No	No	No	No	No	No
N	25892	25812	25884	25633	25891	25892	25892	25888	25892	25428
R-sq	0.859	0.822	0.822	0.891	0.827	0.825	0.884	0.856	0.857	0.843

Note: p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The numbers in parentheses represent standard errors, adjusted for heteroskedasticity and clustered at the state level.

4. Extension on Human population

4.1. Extension on Human population

Katovich (2024) treats human population as a mediator in the causal relationship between shale well development and bird population. The original paper adopts a two-step regression: first testing whether shale wells affect population, then whether population affects bird counts. As neither result shows significant, the paper concludes population is not a mediator. However, James and Smith (Fleming et al., 2015) argue that there is no robust evidence that shale development has a significant impact on human population. By contrast, studies have shown health risks linked to shale extraction (Michalski and Ficek, 2016; Vidic et al., 2013; Institute of Medicine, 2014), leading to siting restrictions in densely populated areas (Miller, 2021; Haley et al., 2016; Mccartney, 2024; Horwitt, 2013). Since human activity may reduce bird populations, population is a pre-treatment confounder affecting both shale development and bird outcomes.

The equation below includes human population as a control in the baseline continuous specification to reduce selection bias. This model requires a conditional independence assumption: conditional on human population, the shale development is independent of potential bird outcomes. Therefore, the model satisfies a randomized assignment.

$$y_{ct} = \beta_1 T_{ct} + \beta_2 HumanPopulation + X_{ct}'\beta_3 + \gamma_c + \delta_t + \epsilon_{ct}$$

Table 5 presents results of this model. Column (1) shows that a 10% increase in shale activity is associated with a 0.39% decline in the total bird population after controlling for local population levels, significant at 5% confident level. A 10% increase in shale activity is significantly associated with declines in grassland birds (–0.43%), woodland (–0.28%), other habitat (–0.41%), urban (–0.36%), non-urban (–0.37%), resident (–0.31%), short-distance migrants (–0.35%), and long-distance migrants (–0.48%). Although estimate coefficients of population are not significantly in most models, its inclusion help reduce omitted variable bias.

Table 5. Extension: Human Population as a Confounder

	Total	Grassland	Woodland	Wetland	Other Habitat	Urban	Non-Urban	Resident	Short Migration	Long Migration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Operating Shale Well	-0.0396**	-0.0442***	-0.0284**	-0.0300	-0.0421*	-0.0365*	-0.0380**	-0.0319***	-0.0354*	-0.0494**
	(0.019)	(0.002)	(0.045)	(0.176)	(0.059)	(0.054)	(0.021)	(0.006)	(0.053)	(0.045)
Population	-0.0522	-0.0936	0.0343	0.0204	-0.195	0.00940	-0.00283	0.150*	-0.0192	-0.0119
	(0.597)	(0.404)	(0.722)	(0.900)	(0.102)	(0.927)	(0.982)	(0.078)	(0.853)	(0.943)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trends	No	No	No	No	No	No	No	No	No	No
N	25937	25856	25929	25676	25936	25937	25937	25933	25937	25471
R-sq	0.856	0.777	0.812	0.879	0.835	0.834	0.860	0.831	0.857	0.814

Note: p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The numbers in parentheses represent standard errors, adjusted for heteroskedasticity and clustered at the state level.

Table 6. Extension: Interaction between Shale Wells and Population

	Total	Grassland	Woodland	Wetland	Other Habitat	Urban	Non-Urban	Resident	Short Migration	Long Migration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Operating Shale Well	-0.108	-0.0626	-0.161*	0.0279	-0.122	-0.117	-0.0385	-0.0398	-0.129	0.0906
	(0.517)	(0.496)	(0.093)	(0.893)	(0.600)	(0.540)	(0.747)	(0.603)	(0.484)	(0.591)
Population	-0.0569	-0.0949	0.0251	0.0261	-0.201	0.00383	-0.00286	0.150*	-0.0257	-0.000624
	(0.572)	(0.400)	(0.793)	(0.875)	(0.103)	(0.971)	(0.982)	(0.079)	(0.809)	(0.997)
Shale Well × Population	0.00608	0.00163	0.0117	-0.00511	0.00705	0.00715	0.0000401	0.000692	0.00831	-0.0124
	(0.658)	(0.832)	(0.133)	(0.764)	(0.714)	(0.649)	(0.997)	(0.914)	(0.583)	(0.377)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Circle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trends	No	No	No	No	No	No	No	No	No	No
N	25937	25856	25929	25676	25936	25937	25937	25933	25937	25471
R-sq	0.856	0.777	0.812	0.879	0.835	0.834	0.860	0.831	0.857	0.814

Note: p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The numbers in parentheses represent standard errors, adjusted for heteroskedasticity and clustered at the state level.

4.2. Heterogeneous effects of Human population

To explore heterogeneous treatment effects, I extend the model by including an interaction term between operating shale wells and human population, allowing for exploring different shale well effects in different human population regions:

$$y_{ct} = \beta_1 T_{ct} + \beta_2 HumanPopulation + \beta_3 T_{ct} \times HumanPopulation + X'_{ct}\beta_4 + \gamma_c + \delta_t + \epsilon_{ct}$$

In this model, β_1 captures the main effect that measures the effect of shale well when there is no human. β_3 is the interaction effect which measures the additional effect of shale well for birds with a one-unit increment of human population.

Table 6 shows the results where the interaction term is statistically insignificant across all bird groups, indicating that population density does not systematically change the effect of shale wells.

5. Conclusion and discussion

In the binary treatment model, shale well development has no statistically significant effect on total bird populations. However, in the continuous treatment model, the effect becomes significant, suggesting that the intensity of shale development matters. This result remains robust under weighted least squares (WLS), which accounts for variation in observation effort, further confirming the significance of the effect. To make causal inference, I control for human population as a potential confounder. The effect of shale wells on bird populations remains statistically significant. Furthermore, interaction analyses provide no evidence that the impact of shale development varies systematically with human population density.

However, these models may suffer from measurement error in the dependent variable, and potential reverse causality. For example, bird population might influence the shale well locations. An instrumental variable approach would be ideal to address both concerns. However, due to data limitations and the absence of an instrument that satisfies the relevance, independence, and exclusion restrictions, I am unable to implement this method.

Word count: 2436

Reference:

Bellemare, M.F., Wichman, C.J. (2020) 'Elasticities and the inverse hyperbolic sine transformation'. *Oxford Bulletin of Economics and Statistics*, 82(1), pp.50–61. Available at: https://onlinelibrary.wiley.com/doi/full/10.1111/obes.12325 (Accessed: 1 May 2025).

Fleming, D. A., Komarek, T. M., Partridge, M. D., & Measham, T. G. (2015). 'The booming socioeconomic impacts of shale: A review of findings and methods in the empirical literature', *University Library of Munich, Germany*. Available at: https://mpra.ub.uni-muenchen.de/68487/ (Accessed: 1 May 2025).

Haley, M. et al. (2016). 'Adequacy of current state setbacks for directional high-volume hydraulic fracturing in the Marcellus, Barnett, and Niobrara shale plays', *Environ Health Perspect.*, 124(9), pp. 1323–1333. Available at: https://doi.org/10.1289/ehp.1510547 (Accessed: 1 May 2025).

Horwitt, D. (2013). ‘Fracking harms won’t respect town boundaries’. *Environmental Working Group (EWG)*. Available at: https://www.ewg.org/news-insights/news/fracking-harms-wont-respect-town-boundaries?utm_source=chatgpt.com (Accessed: 1 May 2025).

Institute of Medicine (2014). ‘Health Impact Assessment of Shale Gas Extraction: Workshop Summary’. Washington, DC: The National Academies Press. Available at: https://doi.org/10.17226/18376 (Accessed: 1 May 2025).

Katovich, E. (2023). ‘Quantifying the effects of energy infrastructure on bird populations and biodiversity’. *Environmental Science & Technology*, 58(1), pp.323-332. Available at: https://pubs.acs.org/doi/abs/10.1021/acs.est.3c03899 (Accessed: 1 May 2025).

Mccartney, G. (2024). ‘New Mexico studies oil drilling restrictions that would hit output, revenue’. *Reuters*. Available at: https://www.reuters.com/business/energy/new-mexico-studies-oil-drilling-restrictions-that-would-hit-output-revenue-2024-10-25/?utm_source=chatgpt.com (Accessed: 1 May 2025).

Michalski, R. and Ficek, A. (2016). ‘Environmental pollution by chemical substances used in the shale gas extraction—a review’, *Desalination and Water Treatment*, 57(3), pp.1336–1343. Available at: https://doi.org/10.1080/19443994.2015.1017331 (Accessed: 1 May 2025).

Miller, A. (2021). ‘Pennsylvania setback regulations for fracking do not prevent setback incidents’. *Harvard T.H. Chan School of Public Health*. Available at: https://hsph.harvard.edu/climate-health-c-change/news/pennsylvania-setback-regulations-for-fracking-do-not-prevent-setback-incidents/?utm_source=chatgpt.com (Accessed: 1 May 2025).

OpenAI (2025) *ChatGPT response to author’s prompt regarding Stata code*, 27 April. Available at: https://openai.com/chatgpt (Accessed: 1 May 2025).

Vidic, R.D. et al. (2013). ‘Impact of shale gas development on regional water quality’, *Science*, 340(6134), pp.1235009. Available at: https://doi.org/10.1126/science.1235009 (Accessed: 1 May 2025).