1 Introduction

We have anecdotal evidence that the Henry-Waha price gap is a good proxy for pipeline congestion, and suggestive evidence (Figure 1) that this price gap is related to emissions. We want to formalize this story, which will involve measuring pipeline congestion and then translating congestion into costs for producers.

2 Model

We propose a simple model of firm behavior. Oil and gas are co-produced. Production is costly, as is gas capture and transmission. Firms choose production intensity and how much of their gas to capture based on costs and commodity prices.

State variables: p^o, p^g (prices), u (pipeline utilization rate)

Control variables: q^o (oil production), e (emissions rate)

$$\max_{q^o, e\in[0,1]} \underbrace{p^o q^o - c^o(q^o)}_{\text{oil profit}} + \underbrace{p^g[(1-e)q^g(q^o)]}_{\text{gas revenue}} - \underbrace{c_a(1-e)q^g(q^o)}_{\text{capture cost}} - \underbrace{c_t(u)[(1-e)q^g(q^o)]}_{\text{transmission cost}}$$

$$\text{FOC:} \quad \underbrace{p^g - c_t(u)}_{\text{shadow price}} = \underbrace{c_a'(1-e)}_{\text{MACC}}$$

Our goal is to use a flexible function f to estimate inverse marginal abatement cost curve (MACC):

$$e = f(p^g - c_t) + \varepsilon$$

We can also use a flexible function g to estimate transport costs:

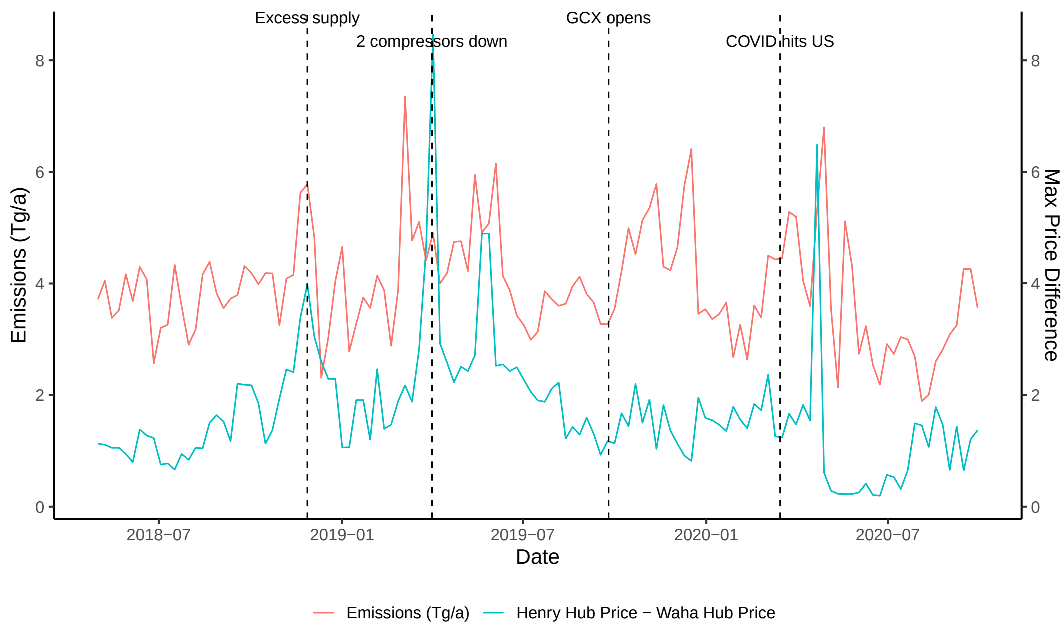
$$c_t = g(u)$$

3 What does the price gap mean?

Based on industry reports around periods of large price gaps, we've concluded that excess supply, infrastructure, and market disruptions all play a role in driving the price gap (Figure 1). We've also thought about why the price gap might be relevant for producers beyond what they learn from the Waha price alone. Here are some of our ideas:

1. Not all Permian gas is sold at Waha Hub.
   - Producers and marketers contract directly with buyers, and these sales will not affect Waha Hub prices.
   - The gap could reflect congestion that is relevant even for non-Waha Hub transactions.
2. Hedging.
   - Producers do not want to be fully exposed to Waha fluctuations.
   - They can hedge by contracting based on Henry Price, in which case they will care about transmission costs (which the price gap proxies) but not the Waha price itself.
3. Congestion between Henry-Waha might correlate with congestion within the Permian.
   - Difficulty moving gas out of the basin could cause backups on local pipelines.
   - If Permian producers are unable to move gas to Waha Hub, the Waha price might not fall enough to clear the market.

Figure 1: Emissions and the Henry-Waha Price Gap



4 How can we measure capacity/costs?

4.1 Receipt/delivery points

We observe daily flows and utilization rates for interstate pipeline receipt/delivery points within the Permian Basin (Figure 2). Each point represents a place in the pipeline network where gas enters the system, is transferred to another pipeline, or exits the system. Points are grouped into categories including gathering, interconnect, processing, compressor, and delivery to a local distribution company.

We could use point-level data to estimate pipeline utilization rates and then calculate transmission costs to be some function of this utilization rate. We could use the Henry-Waha price gap to discipline our cost estimates.

There are two main issues with this approach:

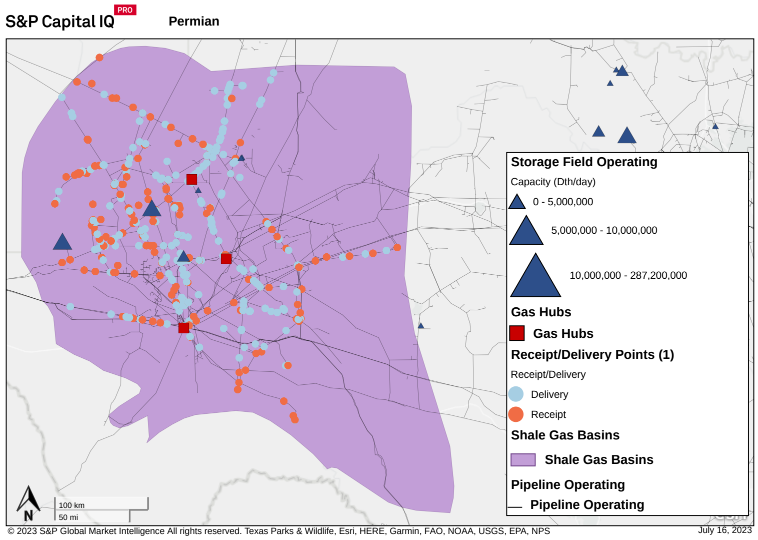
1. **Incomplete data:** We do not observe utilization rates for points along the transmission pipelines linking the Permian to the Gulf Coast because these pipelines are intrastate, and thus subject to more lax reporting requirements.

2. **Inaccurate max capacity numbers:** Pipelines are responsible for reporting on their own flows and capacities. They do not always update available capacity numbers to account for equipment failures and other outages. So, we are unable to confidently distinguish between low demand for pipeline capacity and low available capacity.

After exploring the receipt/delivery point data, there does not appear to be a relationship between flows and the price gap. Flows may decrease due to slowing supply or due to pipeline transportation issues like maintenance events. In order to disentangle changes in flows from supply changes and pipeline disruptions, we can use data on pipeline disruption notices.

We are in the process of purchasing a dataset from Wood Mackenzie on pipeline operator-reported maintenance and outage events. The dataset will include the date of each event, where it took place, and an indicator of its severity. Using this data, we can estimate the utilization of *currently available* capacity.

Figure 2: Permian receipt and delivery points



4.2 Instrument for the price gap

It could be that the price gap is the closest thing we have to a measurement of transmission costs. In that case, we will want to find a source of exogenous variation in capacity, so we can isolate the link between capacity and costs.

One option for an instrument is pipeline maintenance issues from the WoodMackenzie data. We can use this data to instrument for the Henry-Waha price gap, isolating the portion of the gap that is due to exogenous shocks to capacity. Agerton et al. (unpublished) use this approach, instrumenting for the price gap using the number of critical maintenance events in Texas on any given day, controlling for drilling and production. They find that this instrument has a significant and positive first stage coefficient.

Another option is using an instrument capturing exogenous shocks to the oil price (Section 5). Drilling responds to oil prices (Figure 4), and production from existing wells is declining over time (Figure 3). Thus, an exogenous increase in oil prices should induce more new wells, resulting in sharply more gas that needs to be transported via pipeline. In theory, we should be able to link exogenous increases in gas volumes produced to observed changes in the price gap.

5 Oil prices → drilling → emissions?

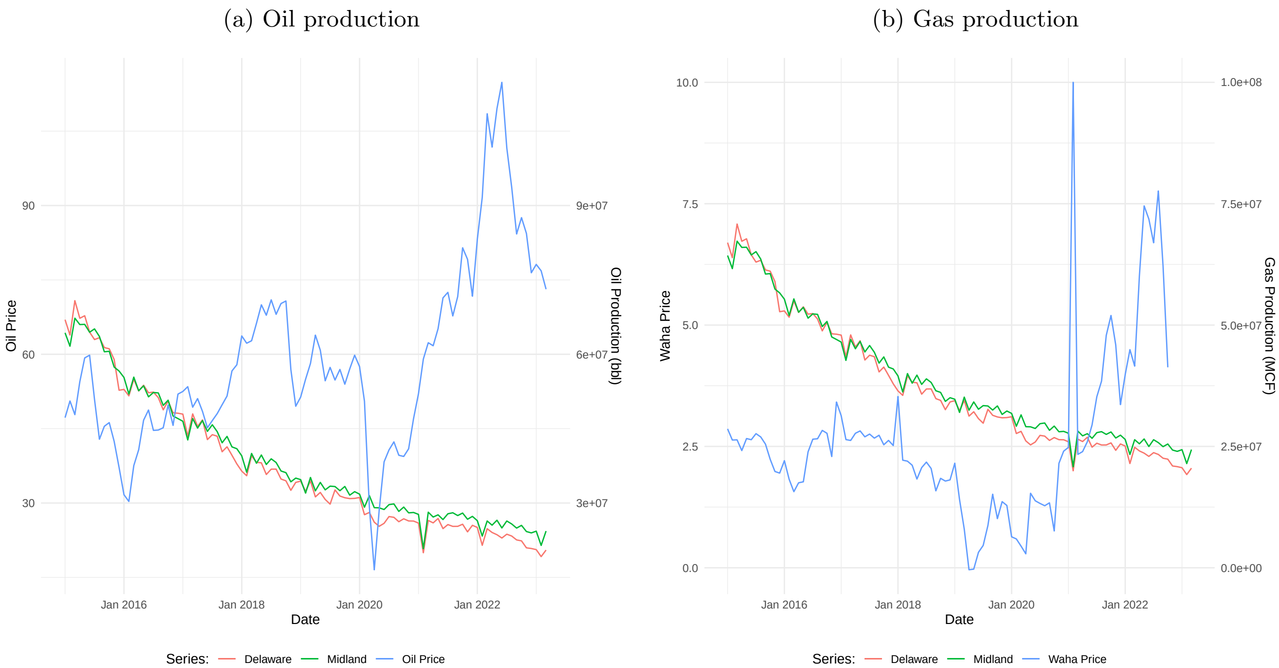
We want to show that higher oil prices induce more drilling, which leads to more gas being both sold and emitted.

5.1 Oil prices and production

Anderson, Kellogg, and Salant (2018) show that oil production responds to oil prices along the extensive margin (drilling new wells), but not the intensive margin (production from existing wells). We present evidence that this pattern holds true in our data. Price fluctuations affect neither gas production nor oil production (Figure 3) from existing wells. Co-movement appears limited to the winter, when freezing weather independently affects both supply (due to frozen equipment) and demand for gas.

As expected, drilling activity ("spudding", or the initial drilling phase for a well) appears to be strongly related to oil prices (Figure 4b). However, it is not responsive to gas prices (Figure 4a). This appears true in both the Delaware and the Midland basins, even though the Delaware Basin has a higher gas-to-oil ratio.

Figure 3: Production from existing wells



Notes: We restrict the sample here to wells that existed as of January 2015.

Figure 4: Well drilling and prices

When we regress drilling (spud date) and well completions (first production date) against prices, the same pattern emerges (Table 1). Drilling and completions both increase significantly when oil prices are higher. Both also have a negative relationship with 3 month oil futures, perhaps because producers anticipate lower prices in the future by re-timing production towards the present. Drilling and completions are much less responsive to benchmark natural gas prices. Production from existing wells (column 3) has a nearly significant but negative correlation with oil prices, which could be an artifact of winter supply/demand patterns.

Table 1: Spuds, New Wells, Production from Existing Wells

	Dependent variable:		
	log(sum_spuds) Spuds	log(Firstprod) First Prod	log(SUM.Monthly.Oil.) Bbl Oil
	(1)	(2)	(3)
Log(Cushing, Forward contract)	2.711^{***}	2.564^{***}	-1.731^{*}
	(0.810)	(0.534)	(0.881)
Log(Cushing, 3 month contract)	-2.443^{***}	-2.423^{***}	1.778
	(0.861)	(0.567)	(1.124)
Log(Henry spot price)	−0.226	-0.190^{*}	−0.210
	(0.163)	(0.109)	(0.194)
Basin = Midland	0.333^{***}	0.643^{***}	-1.147^{***}
	(0.081)	(0.053)	(0.087)
Production type = Oil			3.673^{***}
			(0.087)
Constant	4.105^{***}	4.414^{***}	12.527^{***}
	(0.601)	(0.395)	(1.319)
Observations	320	318	396
R^2	0.099	0.358	0.836
Adjusted R^2	0.088	0.350	0.834
Residual Std. Error	0.721 (df = 315)	0.474 (df = 313)	0.862 (df = 390)
F Statistic	8.665^{***} (df = 4; 315)	43.678^{***} (df = 4; 313)	396.896^{***} (df = 5; 390)

Notes: An observation is a basin-month. Prices are in $2018 and reflect the average of prices over the month. Oil production (column 3) is restricted to wells that existed prior 2015, and so the data period is 2015-2023. For other specifications, the data period is 2010-2023.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

5.2 Production and emissions

For the second link in our causal chain from oil prices to emissions, we want to establish how much emissions come from drilling and production volume, outside of producer decisions to vent or flare. Knowing this, we can determine the increase in emissions induced by producer decisions to flare/vent because of economic factors (e.g. limited transport capacity due to more total gas production).

There are several reasons why increased drilling and production would increase emissions even if all gas that can be sold is sold. The process of completing a well typically involves venting gas. It is also common for there to be some lag time between well completion and pipeline connection; gas produced in the meantime must be vented or flared. Finally, it is reasonable to assume that some share of gas produced is unintentionally vented due to pipeline leaks and imperfect equipment, and so more production should mechanically induce more emissions.

We examine the relationship between drilling/production and emissions using county-month level data on methane emissions and production in Table 2. We cannot interpret these results causally, since production might respond to gas pipeline constraints, which themselves might influence emissions. Regardless, the coefficients we estimate have the expected signs: drilling, completions, and gas production are all correlated with more emissions. We hope to do more work to

Table 2: Spuds, New Wells, Production from Existing Wells

	Dependent variable:			
	emissions_ggm			
	(1)	(2)	(3)	(4)
Spuds	3.477^{***}			1.164^{***}
	(0.202)			(0.287)
Completions		4.122^{***}		1.121^{***}
		(0.215)		(0.334)
Gas production			0.00001^{***}	0.00000^{***}
			(0.00000)	(0.00000)
Basin = Delaware	110.664^{***}	90.279^{***}	54.893^{***}	3.591
	(13.752)	(14.154)	(11.179)	(14.697)
Basin = Midland	4.285	-7.010	-14.900^{**}	-20.528^{**}
	(10.079)	(10.311)	(6.358)	(8.463)
Basin = Val Verde	54.100^{***}	56.278^{***}	-13.368	29.223^{*}
	(19.001)	(18.988)	(9.233)	(15.107)
Constant	38.784^{***}	33.577^{***}	42.970^{***}	30.978^{***}
	(7.979)	(8.744)	(5.725)	(6.944)
Observations	522	522	870	464
R^2	0.526	0.556	0.725	0.740
Adjusted R^2	0.523	0.553	0.723	0.736
Residual Std. Error	93.004 (df = 517)	91.009 (df = 517)	68.029 (df = 865)	71.964 (df = 457)
F Statistic	143.661^{***} (df = 4; 517)	161.894^{***} (df = 4; 517)	569.376^{***} (df = 4; 865)	216.408^{***} (df = 6; 457)

Notes: An observation is a county-month. The data period is May 2018-Sep 2020. Emissions are measured in gigagrams/year.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

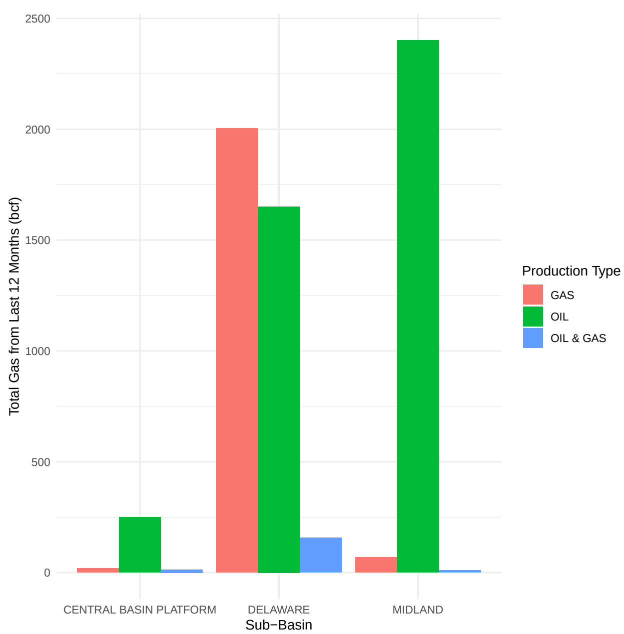
6 To-dos

- Gather more industry background on pipeline engineering, bottlenecks
- Generate rough estimate for share of emissions due to drilling and new wells vs. flaring/venting decisions
- Figure out how to model costs
- Find instrument for oil price changes

A Appendix: Figures and Tables

Figure 5: Production by basin

(a) Gas production



(b) Oil production